



الاختبارات & الامتحانات

قياس القدرات والأداء

المحررة: لويس ر. أيكين



المحررة
لويس ر. أيكين

الاختبارات والامتحانات
قياس القدرات والأداء

مكتبة العبيكان

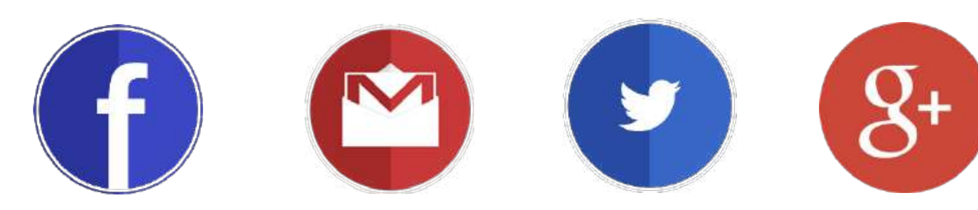








mohamed khatab




الاختبارات والامتحانات
قياس القدرات والأداء

الاختبارات والامتحانات
قياس القدرات والأداء

المحررة:

لويس ر. أيكين

نقلته إلى العربية

د. فرح السراج

راجعه

د. عبدالمطلب يوسف جابر



Original Title:

Tests & Examinations

Measuring Abilities

and Performance

By: Lewis R. Aiken

ISBN 0 - 471 - 19263 - 5

ISBN 5 - 149 - 54 - 9960

الطبعة العربية الأولى 1428هـ - 2007م

الناشر

شركة العبيكان Obeikan للأبحاث والتطوير

المملكة العربية السعودية - شارع العليا العام - جنوب برج المملكة - عمارة الموسى للمكاتب

هاتف: ٢٩٣٧٥٧٤ / ٢٩٣٧٥٨١، فاكس: ٢٩٣٧٥٨٨ ص. ب: 67622 الرياض 11517

© مكتبة العبيكان، 1427هـ

فهرسة مكتبة الملك فهد الوطنية أثناء النشر

رايكين، لويس

الاختبارات والامتحانات قياس القدرات والأداء. لويس رايكين؛ فرح السراج. - الرياض 1428هـ

446 ص؛ 16,5× 24 سم

ردمك: 5 - 149 - 54 - 9960

1 ـ الاختبارات والمقاييس التربوية أ. السراج، فرح (مترجم) ب. العنوان

ديوي: 371,26 1427 / 6929

ردمك: 5- 149 - 54 - 9960 رقم الإيداع: 6929 / 1427

هذا الكتاب من كتب مشروع الترجمة المشترك بين وزارة التعليم العالي وشركة مكتبة العبيكان

امتياز التوزيع شركة مكتبة العبيكان Obeikan

المملكة العربية السعودية - العليا - تقاطع طريق الملك فهد مع شارع العروبة

هاتف: ٤١٦٠٠١٨ / ٤٦٥٤٤٢٤ - فاكس: ٤٦٥٠١٢٩ ص. ب: 62807 الرياض 11595



تقديم معالي وزير التعليم العالي

الحمد لله والصلاة والسلام على رسول الله وبعد:

تحرص وزارة التعليم العالي في المملكة العربية السعودية على تشييد بنية متينة للتعليم العالي في المملكة تأخذ في الحسبان متطلبات مجتمعها وثقافته الإسلامية العريقة، وفي الوقت نفسه تحاكي أنظمة التعليم العالي العالمية. وكان الغرض الأساس للسعي وراء هذا الهدف هو تطوير العملية التعليمية، وكذلك تطوير النظام الإداري المصاحب خاصة في ضوء الطفرة المعلوماتية والعولمة والمنافسة الشديدة بين مؤسسات التعليم العالي على المستويات المحلية والإقليمية والدولية.

ونظراً لما حققه التعليم العالي في المملكة العربية السعودية من تطور كمّي ونوعي بدعم سخي من حكومتنا الرشيدة بقيادة خادم الحرمين الشريفين الملك عبدالله بن عبدالعزيز، وسمو ولي عهده الأمين، الأمير سلطان بن عبدالعزيز -يحفظهما الله- فقد ظهرت الحاجة بشكل أكبر لتوفير المصادر المختلفة لتعزيز توعية الأفراد العاملين في حقل التعليم العالي بما ينشر في هذا المجال باللغات الأجنبية. لذا، رأت وزارة التعليم العالي ترجمة عدد من الكتب ذات العلاقة بمجالات التطوير الأكاديمي وتقديمها باللغة العربية لتكون في متناول جميع العاملين في القطاع الأكاديمي. ونظرا لقلة مثل هذه الكتب في المكتبة العربية، فقد سعت الوزارة إلى توفيرها بشكل سريع وفاعل، وعليه كان مشروع الترجمة هذا. ولقد قامت الوزارة باختيار كتب تحوي دراسات حازت قبولا وانتشارا في الكثير من المؤسسات التعليمية ذات الشهرة العالمية وأنجزت بأيدي عدد من

الأكاديميين والإداريين المهتمين بالتطوير في التعليم العالي. وعالجت الدراسات في هذه الكتب قضايا متعلقة بكل من تطوير مهارات الأساتذة ورؤساء الوحدات الأكاديمية والإداريين في أكثر الجامعات العالمية تقدما. كما تناولت هذه الكتب قضايا مثل: التعليم الإلكتروني، والتعليم عن بعد، ومهارات التعليم والتعلم، وتقنيات التعليم الحديثة، والتخطيط الاستراتيجي الخاص بالتعليم، والاختبارات والتقويم ,ومواءمة مخرجات التعليم العالي لسوق العمل، وتحقيق الجودة في مدخلات ومخرجات التعليم العالي وغير ذلك من الموضوعات ذات العلاقة.

ووقع اختيار الوزارة على مكتبة العبيكان للنشر للتعاون معها في نشر ترجمات هذه السلسلة من الكتب الأكاديمية المتخصصة وذلك لما لهذه المكتبة من خبرة وتميز في مجال النشر وفي ميداني التأليف والترجمة والكفاءة في الأداء. وقامت مكتبة العبيكان بمهمة الاتفاق مع الناشرين للكتب الأجنبية ومن ثم ترجمتها وتقديمها للقارئ بالشكل المناسب، وقد تم مراجعة هذه الكتب من قبل فرق أكاديمية متخصصة.

وتأمل الوزارة بأن تكون بهذا المشروع قد أسهمت بوضع دليل متكامل من الدراسات المهمة والمشروعات والأفكار ذات العلاقة بتطوير التعليم العالي بين أيدي جميع أعضاء الهيكل الأكاديمي و الإداري في الجامعات ابتداء من مديري الجامعات إلى أول الصاعدين على سلم التعليم والإدارة فيها.

وإذ تقدم هذه الكتب وأفكارها خلاصة تجارب المجتمعات الأكاديمية المتطورة في هذا المجال فإنها لا تقلّل من الخبرات ولا التجارب الميدانية المحلية لدينا، وتلك المستمدة من ديننا الحنيف وثقافتنا بل إنها ستعزز دور المجتمع الأكاديمي والإسهام في بناء وطننا الكريم، كما ستساعدنا على التخلص من الأخطاء التي مررنا بها أو وقعت لغيرنا فنتجنب تكرارها.

ولا يفوتني أن أشكر معالي الدكتور خالد بن صالح السلطان مدير جامعة الملك فهد للبترول والمعادن، وسعادة الدكتور سهل بن نشأت عبدالجواد، عميد التطوير الأكاديمي في الجامعة، وجميع من عمل معهم على جهودهم المباركة لإخراج هذا المشروع إلى أن أصبح واقعا ملموساً وجهداً متميزاً، والذي سيكون له -بإذن الله- مردود إيجابي على المجتمع.

وفي الختام يسرنا أن تنشر وزارة التعليم العالي في المملكة العربية السعودية بالتعاون مع مكتبة العبيكان للنشر هذه السلسلة من ترجمات الكتب الأكاديمية المتخصصة، ونأمل أن تكون دليلا معرفيا يسهم في التطوير والتنمية، وذلك بجانب ما توافر في السابق لننطلق للمستقبل بأحسن ما توافر لدينا من خبراتنا الخاصة وما نتعلمه من تجارب الآخرين في جوانب البحث العلمي والأكاديمي في العالم... والله ولي الموفق،،،

الدكتور خالد بن محمد العنقري

وزير التعليم العالي في المملكة العربية السعودية

المحتويات

الموضوع	الصفحة

المقدمة

يضم هذا الكتاب معالجة مختصرة، إلا أنها شاملة، لعملية التقييم النفسي والتعليمي في المجال الإدراكي. ويستهدف الكتاب بشكل أساسي الباحثين والعاملين في مجال العلوم السلوكية من الذين يستخدمون الاختبارات الموضوعية والاختبارات المقالية والاختبارات الشفهية واختبارات الأداء في أبحاث نظرية وتطبيقية تدرس عمليات إنتقاء (الطلاب والموظفين)، وتنسيب وترقية (الموظفين فقط)، ومنح الشهادات والتراخيص، وتقييم آثار العملية التعليمية أو البرنامج التدريبي أو الإجراءات العلاجية على السلوك والإدراك.

يستهدف هذا الكتاب جميع العاملين في المجالات التي تتطلب إعداد وإجراء اختبارات يضعها العاملون بأنفسهم أو إجراء اختبارات مقننة لخدمة أغراض نظرية أو تطبيقية كمسح وإنتقاء وتنسيب الطلاب أو الموظفين، وتقييم الأداء، وتقييم المقررات والبرامج التعليمية والتدريبية، والتخطيط التنظيمي. ومن ضمن المستفيدين من هذا الكتاب مسؤولو شؤون الموظفين والمدرسون والإداريون وغيرهم من العاملين في المؤسسات التي تعنى بدراسة القدرات الإدراكية وتطويرها عند الموظفين والطلاب وغيرهم من زبائن المؤسسة.

ومن الممكن كذلك استخدام الكتاب كمرجع لمقررات البحث العلمي في مجالات الإقتصاد وعلم النفس والتربية وعلم الاجتماع والطب أو لورشات العمل المعنية بالتقييم التعليمي والنفسي. كما سيجد طلاب الدراسات العليا الذين هم في صدد كتابة أطروحات الماجستير أو الدكتوراة في صفحات هذا الكتاب العديد من الاقتراحات البناءة في مجال تقييم السلوك والتحصيل البشري.

هذا الكتاب هو الجزء الثالث من سلسلة من الكتب حول المقاييس النفسية، صدر الجزء الأول والثاني منها تحت عنواني: سلالم التقدير وقوائم الرصد (أيكين، ١٩٩٦) والاستبانات ومقاييس السمات (أيكين، b ١٩٩٧). وعلى الرغم من أن الكتابين المذكورين آنفاً يتناولان موضوع تقييم القدرات البشرية، لا يتطرق هذا الكتابان بشكل خاص إلى التقييم الإدراكي واختبارات التحصيل والقدرات الخاصة بل يركزان على الصفات العاطفية عند الأفراد كسمات الشخصية والطباع والآراء. يدرس هذا الكتاب القدرات الحركية – الحسية بشكل محدود ويستفيض في دراسة المقاييس الإدراكية مما يختم الدراسة الشاملة لسلسلة المقاييس النفسية.

يتألف هذا الكتاب (شأنه كشأن المجلدين السابقين من كتب السلسلة) من ثمانية فصول وعدد من الملاحق. كما يُرفق بالكتاب قرص مرن يضم عدداً كبيراً من برامج الحاسب المعنية بتصميم وإجراء وتصحيح وقياس وتقييم الاختبارات وغيرها من وسائل تقييم الأداء والمشاريع والمنتجات. وكما هو الحال في المجلدين السابقين يمكن استخدام الكتاب والبرامج مجتمعةً أو كلاً على حدة. كذلك يضم هذا الكتاب معلومات نظرية ومفاهيمية وتطبيقات عملية لمختلف أنواع الاختبارات التي تستخدم في شتى أنواع المواقف والحالات. يعالج الكتاب وبالتفصيل الاختبارات الكتابية والشفهية واختبارات الأداء والتحصيل واختبارات القدرات الإدراكية العامة (الذكاء) واختبارات القدرات والاستعدادات الخاصة. كما يفيد فهرس المصطلحات والملاحق المختلفة للكتاب وقوائم الكتب المقترحة في نهاية كل فصل في فهم المفاهيم والمبادئ والإجراءات التي تم التحدث عنها في فصول الكتاب وفي التوجه نحو مصادر أبعد مما تم تقديمه في هذه الفصول.

لويس ر. أيكين

١

الجذور والقضايا المهنية

تطالعنا المعاجم اللغوية بتعاريف عديدة لكلمة "اختبار" إلا أن المعنى الذي سيستخدم في هذا الكتاب هو معنى الاختبار النفسي ويعني مجموعة الأسئلة والمسائل والمهام المصممة لاستثارة ردود وإجابات يتم استخدامها لاحقاً في تحديد وقياس صفات الأفراد أو إمكاناتهم أو تحصيلهم. وغالباً ما تكون الاختبارات التي تحدد صفات الأشخاص اختباراتٍ للشخصية أو مقاييس للسمات أو مؤشرات للقياس، وقد تم تناول موضوع اختبارات الشخصية بشكلٍ تفصيلي في كتابين سابقين من هذه السلسلة هما: سلالم التقدير وقوائم الرصد، والاستبانات ومقاييس السمات. وغالباً ما ينظر إلى هذا النوع من أدوات اختبار الشخصية على أنها مقاييس عاطفية لاإدراكية أو لاذهنية تقيم دوافع واهتمامات ومواقف وميول الأفراد بالإضافة إلى خصائص وصفات أساليبهم الشخصية. يتناول هذا الكتاب، على عكس الكتب الآنفة الذكر، أدوات قياس الشخصية المعنية بالقدرات الإدراكية للأفراد والتي تتألف من اختبارات تقيس عمليات ومنتجات الذهن من ذاكرة واستيعاب واستدلال وحل للمشكلات.

وعلى الرغم من أن معظم هذه الاختبارات تتطلب استخدام ورقة وقلم إلا أن بعضها يمكن أن يكون شفهياً أو اختباراً للأداء. وتقسم الاختبارات الإدراكية إلى اختباراتٍ تحصيلية (Achievement Tests) تقيس درجة معرفة الفرد أو مهارته ضمن مادة علمية معينة أو مشروع إنتاجي معين وعند النقطة الزمنية التي يتم فيها الإختبار، واختبارات للاستعداد (Aptitude Tests) تقيس قدرة الأفراد على تعلم مهارة أو مادة علمية معينة. ويكمن الاختلاف الرئيسي بين اختبارات الاستعداد والاختبارات التحصيلية في كون الأخيرة أدوات تقيس الماضي والحاضر أو ما تمكن الفرد من تعلمه وتحصيله بينما تعنى اختبارات الاستعداد بقياس الإمكانيات المستقبلية للفرد أو ما يمكن للشخص تعلمه في حال توفر التعليم والتدريب المناسبين له. ونظراً لأن كلمة "قدرة" تشمل مفاهيم الذكاء والتحصيل والاستعداد، يشار عادةً إلى هذا النوع من المقاييس الإدراكية على أنها اختبارات قدرة، وعليه تعتبر اختبارات الذكاء اختبارات للقدرات العقلية بشكلٍ عام بينما يقتصر عمل اختبارات الاستعداد على قياس إمكانيات الأفراد في مجالات محددة فقط.

مما لا شك فيه أن كلمة "اختبار" كلمة مكروهة من قبل الملايين من الناس الذين يتعين عليهم كل عام أن يثبتوا درجة معرفتهم ومهارتهم في مادة دراسية ما أو في حرفة معينة أو في غيرهما من المجالات التي يتطلب نجاح أو استمرار الأفراد فيها تحقيقهم لمعاييرأداء معينة. ويرى الناس الامتحانات والاختبارات كعقبات وعوائق تحول بينهم وبين الحصول على الإنتماء الكامل لطبقة معينةأو مؤسسة أو مهنة أو أي مجموعة تخصصية يرغبون في الإنتماء إليها. ومن الجدير بالذكر أن هنالك امتحانات لا يترتب عليها إلتحاق شخصٍ ما (أو عدم إلتحاقه) ببرنامج تعليمي أو تدريبي أو مهني معين كامتحانات القدرة التي تهدف

إلى تصنيف الأشخاص ضمن البرنامج أو المنصب أو مجال العمل الأنسب لهم. وتهدف أنواع أخرى من الإختبارات إلى تحديد مصير بعض الأشخاص من حيث منحهم شهادة أو وثيقة تخرج أو غيرهما مما يوثق اجتياز هؤلاء الأشخاص وبنجاح مستوى الأداء المقبول. كما تستخدم أنواع أخرى من الإختبارات كأدوات تشخيصية لتحديد المشكلات أو الاحتياجات الخاصة عند بعض الأفراد أو المجموعات ومن ثم وضع خطط لمعالجتها.

الخلفية التاريخية

لقد قاس الإنسان ومنذ فجر التاريخ القدرات البشرية بطريقةٍ أو بأخرى وعلى الرغم من أن هذه "الاختبارات" لم تكن علميةً بمقاييس اليوم إلا أنها وبدون شك كانت على درجةٍ من الشرعية سمحت باستخدامها حقباً طويلة من الزمن. إن مجرد مراقبة أشخاصٍ يؤدون مهمةً ما يقودنا إلى ملاحظة درجات المهارة المتفاوتة التي ينفذ بها هؤلاء الناس المهام المسندة إليهم وقدرة البعض منهم على إداء المهمة بنجاح في حال فشل الآخرين في الإرتقاء إلى المستوى المطلوب، ويقودنا هذا التصنيف بدوره إلى ترتيب مؤديّ المهمة من الأكثر فشلاً إلى الأكثر نجاحاً وبالتالي إلى وضع نظامٍ لتقييم اختبارات الأداء. كما يمكننا زيادة درجة دقة وتناغم عملية التقييم هذه من خلال إخضاع جميع المختَبرين إلى ظروف الاختبار ذاتها من حيث المواد المستخدمة والموقع والتوقيت ووجود أشخاص آخرين، وبهذه الطريقة يتم وضع امتحان يشكل المقياس أو المعيار الذي يتم على أساسه تقييم أداء الأفراد للمهمة التي يتألف منها الاختبار. ولعل أول إختبارغير رسمي في تاريخ البشرية هو "اختبار التفاحة" المذكور في سفر التكوين من الإنجيل. وفي هذا الاختبار لم يكن هناك إلا ممتَحَنين هما آدم وحواء إلا أن عواقب فشلهما كانت وخيمة. ويذكر الإنجيل

مثالاً آخر عن الامتحانات إذ يخاطب الله شعيباً في سفر القضاة ٧-٢:٧ بقوله:

> وقال الرب لجدعون: إن لديك أناساً كثيرين لأدفع مدينا في قبضتهم وذلك حتى لا تفاخر عليَّ إسرائيل بأنها خلصت نفسها بقوتها، والآن أعلن للناس بأنه من كان خائفاً ومرتعداً فليرجع وينصرف من جبل جلعاد. فرجع من الشعب اثنان وعشرون ألفا وبقي عشرة آلاف.
>
> وقال الرب لجدعون: لم يزل الشعب كثيراً انزل بهم إلى الماء فسأمحصهم لك هناك؛ فالذي أقول لك عنه يذهب معك فهو يذهب معك وكل من أقول لك عنه هذا لا يذهب معك، فهو لا يذهب.
>
> فنزل بالناس إلى الماء وقال الرب لجدعون: قم بفصل كل من لصق الماء بلسانه كما يلعق الكلب عن أولئك الذين يجثون على ركبهم للشرب. وكان عدد الذين لعقوا بيدهم إلى فمهم ثلاث مئة رجل وأما باقي الناس فجثوا على ركبهم لشرب الماء.
>
> قال الرب لجدعون بالثلاث مئة رجل الذين لعقوا بألسنتهم أخلصكم وأدفع المدينيين ليدك، وأما سائر الناس فليذهبوا كل إلى مكانه.

يبين هذا المقطع أن الله يعلم أن الجنود الذين لعقوا الماء كالكلاب هم أكثر حذراً وتأهباً ولذلك فإن عدة مئاتٍ من هؤلاء الجنود أكفأ بكثير من جحافل جنودٍ أقل حذراً وأقل تأهباً.

وتعود جذور الاختبارات النفسية الرسمية إلى الصين عام ٤٠٠٠ قبل الميلاد إذ وضع إمبراطور الصين آنذاك نظاماً امتحاناً يهدف إلى تحديد قدرة الأفراد على استلام مناصب في الدولة من خلال امتحانات شفهية تقام كل ٣ سنوات.

وبعد ٢٠٠٠ عاماً من ذلك الإجراء، قامت أسرة هان الحاكمة بإخضاع موظفي الدولة إلى إختبارات كتابية في القانون المدني والشؤون العسكرية والزراعة والمالية والجغرافية والإنشاء والشعر. وفي عام ١٣٧٠ ميلادية تم إتخاذ الخطوة الأخيرة في تطور الامتحانات الوظيفية في الصين عندما تم (وتبعاً للفلسفة الكونفوشية) تقسيم الامتحانات إلى ثلاثة أقسام تبدأ باختبارات تمهيدية ينتقل الناجحون فيها إلى اختبارات إقليمية ينتقل الناجحون فيها بدورهم إلى مرحلة الامتحانات على مستوى الدولة، وكانت نسبة الناجحين في امتحانات المرحلة الثالثة لا تتجاوز ١٪ من عدد المتقدمين للامتحانات التمهيدية. وقد استوحت عدد من الدول الأوروبية كبريطانيا وفرنسا وألمانيا امتحاناتها الوظيفية من النموذج الصيني خلال القرن التاسع عشر، إلا أن الصينيين أنفسهم تخلوا عن نظام امتحاناتهم مع بداية القرن العشرين.

الجدول ١-١

أحداث مختارة من تاريخ الاختبارات النفسية والتعليمية

٢٢٠٠ ق.م. استخدام الصينيين للامتحانات الوظيفية
١٢١٩ اعتماد الجامعات الأوروبية لنظام الامتحانات الرسمية
١٥٤٠ بدء الامتحانات الكتابية في الجامعات
١٥٩٩ نشر اليسوعيين لقوانين الامتحانات الكتابية
١٨٣٣ بدء نظام الامتحانات التنافسية لوظائف الدولة في بريطانيا
١٨٤٥ استخدام لجنة مدرسة بوسطن لامتحانات مطبوعة لأول مرة تحت إشراف التربوي هوراس مان
١٨٦٤ إعداد مدرس بريطاني يدعى جورج فيشر لعدد من المقاييس مؤلفة من نماذج أسئلة وأجوبة بغرض استخدامها كمقياس لتقييم إجابات الطلاب على أسئلة الامتحانات المقالية

١٨٦٦ نشر أ. إ. سيجوين لأول نص يتناول موضوع التقييم ومعالجة التخلف العقلي

١٨٦٩ بدء عملية دراسة الفروقات بين الأفراد بطريقة علمية بعد نشر كتاب تصنيف البشر وفقاً لملكاتهم الطبيعية لفرانسيس غالتون

١٨٧٠ بدء نظام امتحانات التوظيف التنافسية في الولايات المتحدة

١٨٨٤ افتتاح فرانسيس غالتون لمعرض المختبر الانثروبومتري(١) للصحة الدولية

١٨٨٨افتتاح ج. م. كاتل مختبراً للامتحان في جامعة بنسلفانيا

١٨٩٣عرض جوزيف جاسترو لاختبارات المهارات الحسية - الحركية في المعرض الكولومبي في شيكاغو

١٨٩٧انشر ج. م. رايس لنتائج أبحاثه حول قدرات التهجئة عند طلاب المدارس في الولايات المتحدة

١٩٠٤وضع تشارلز سبيرمان لنظرية ثنائية العامل في القدرة الذهنية

١٩٠٥نشر النسخة الأولى من مقياس بينيه - سيمون للذكاء

١٩٠٨نشر النسخة المراجعة من مقياس بينيه - سيمون للذكاء ونشر اختبارات للحساب من قبل ج. ك. ستون و س. أ. كورتيس

١٩٠٨-١٩١٤ تطوير إ. ل. ثورندايك لاختبارات قياسية في الحساب والخط واللغة والتهجئة

١٩١٥تعميم أول اختبارمنشور متعدد الإختيارات وهو اختبار كنساس للقراءة الصامتة. نشر اختبار ستنكويست للقدرة الميكانيكية

١٩١٦انشر لويس تيرمان لمقياس ستانفورد - بينيه للذكاء.

تطوير آرثر أوتيس لأول اختبار جماعي للذكاء مبني على مراجعة تيرمان ستانفورد لمقياس سيمون - بينيه

١٩١٧ وضع واستخدام اختباري الجيش آلفا وبيتا وهما أول إختبارين جماعيين للذكاء عند المنتسبين للجيش الأمريكي

١٩١٩انشر مقياس سيشور للموهبة الموسيقية

١٩٢٣ نشر الطبعة الأولى من اختبار ستانفورد للتحصيل

(١) فرع من علم الأجناس البشرية يبحث في مقاييس الجسم البشري. (المترجم).

١٩٢٤ نشر امتحان آيوا للمحتوى الخاص بالمدارس الثانوية

١٩٢٦ إجراء أول امتحان للاستعداد الدراسي (SAT)

١٩٢٧ نشر جداول جيسيل الإرتقائية

١٩٣٢ نشر اختبار متروبوليتان للتحصيل

١٩٣٨ نشر أ. ك. بيروس لأول كتاب سنوي للمقاييس العقلية.

نشر امتحان القدرات الذهنية للمرحلة الإبتدائية ومصفوفات ريفين التقدمية وامتحان بندر الحركي – البصري على مبدأ الغشتالت

١٩٣٩ نشر مقياس ويكسلر – بلفيو لذكاء الأطفال

١٩٤١ تطوير وتطبيق امتحان تصنيف الطلاب الضباط في سلاح الجو على طيارين متدربين في سلاح الجو الأمريكي

١٩٤٧ نشر بطارية اختبارات الاستعداد Aptitude Test Battery التبايني

١٩٤٩ نشر مقياس ويكسلرلذكاء الأطفال

١٩٦٨ نشر مقياس ويكسلرلذكاء الأطفال ما دون سن المدرسة وأطفال المرحلة الإبتدائية (WPPSI)

١٩٧٠– اليوم، استخدام الحاسب بشكل متزايد لتصميم وإجراء الامتحانات ثم وضع درجاتها وتحليل وتأويل نتائجها

١٩٧٤انشرالنسخة المراجعة لمقياس ويكسلر لذكاء الأطفال (WISC-R) وبطارية اختبارات هالستد – رايتن النفسي – العصبي

١٩٨٠– اليوم، تطوير نظرية الإستجابة لفقرة الاختبار (IRT)

١٩٨٣ نشر بطارية كوفمان لتقييم الأطفال (K-ABC)

١٩٨٥ نشركتاب معايير الاختبارات النفسية والتعليمية

١٩٨٦ نشر مقياس ستانفورد – بينيه للذكاء (الطبعة الرابعة) (SB-IV)

١٩٩٥ نشر الطبعة ١٢ من الكتاب السنوي للمقاييس العقلية

١٩٩٧ نشر الطبعة الثالثة من مقياس ويكسلر لذكاء البالغين (WAIS-III)

كانت الدول الأوروبية أبطأ من الصين في عملية إجراء امتحانات رسمية كان أولها امتحانات القانون في جامعة بولونيا التي بدأت عام ١٢١٩، وقد كانت معظم هذه الامتحانات شفهية ولم تصبح رسمية إلا في القرن السادس عشر. يوضح الجدول ١-١ بعض الأحداث الهامة في تاريخ الامتحانات النفسية والتعليمية، ولعل من أهم هذه الأحداث أول استخدام لامتحانات مطبوعة في مدارس بوسطن عام ١٨٤٥، ووضع معايير موضوعية لتقييم الإجابات على الأسئلة المقالية عام ١٨٦٤، والعمل الرائد الذي بدأه فرانسيس غالتون عام ١٨٦٩ لتقييم الذكاء البشري والقدرات العقلية.

قياس الذكاء

تعتبر أبحاث وكتابات فرانسيس غالتون عن الذكاء في النصف الثاني من القرن التاسع عشر بداية عصر الاختبارات النفسية المعاصرة. وقد كان غالتون على قناعة بأن الذكاء هو قدرة عامة موروثة يمكن قياسها عن طريق اختبارات حسية – حركية بسيطة لمكونات الذكاء المختلفة. ومن المهم هنا الإشارة إلى أن فهم أبحاث غالتون عن الذكاء يتطلب منا أن نتذكر أن غالتون بدأ أبحاثه في وقت كانت فيه عمليات البحث والتجربة المخبرية تتطور بشكل سريع في مجالات العلوم التطبيقية، إلا أن الفهم الفلسفي لعمل العقل البشري والبحث العلمي في عمل العقل والجهاز العصبي كانا بدائيين للغاية. وقد كانت الفلسفة الغالبة عن العقل آنذاك ترى أن العقل يتطور عن طريق تعقيد وتحليل الإنطباعات الحسية، مما يعني أن العقل يبدأ بالحواس ومما يفسرقرار غالتون بقياس الذكاء عن طريق اختبارات حسية – حركية.

وفي ضوء ذلك؛ فإنه ليس من المستغرب أن غالتون اختار لقياس الذكاء والتميز بشكل خاص وسائل من اختبارات قياس حسية – حركية. وكانت اختبارات غالتون تتألف من اختبارات للسلوك مبنية على التمييز الحسي

واختبارات لأوقات ردود الفعل تجاه مثيرات سمعية وبصرية واختبارات لسرعة الحركة والقوة العضلية والإحساس بالألم والدقة السمعية والبصرية وتمييز الأوزان. كما حصل غالتون على أعداد كبيرة من القياسات الخاصة بكل من الأطوال والأوزان وطول ومحيط الرأس والمسافة بين الذراعين وطول الأصبع الأوسط وطول القسم الأسفل من الذراع وحجم استيعاب الرئتين للهواء وغيرها من الخصائص الجسدية. وقد أفادت النتائج التي خلص إليها كلارك ويسلر في دراسة أجراها عام ١٩٠١ على ما أسماه ج. ماكين كاتل (١٨٩٠) "باختبارات غالتون العقلية" بعدم وجود أية علاقة وثيقة بين مقاييس غالتون بعضها مع بعض ومع نتائج الاختبارات الجامعية. وقد أدى ذلك إلى إيقاف كافة المحاولات لقياس الذكاء عن طريق اختبارات "الأدوات النحاسية" التي كانت متوفرة في المختبرات النفسية آنذاك.

وعلى الرغم من أن فشل اختبارات غالتون وكاتل ثبط همم الباحثين في مجال اختبارات الذكاء البسيطة والنظريات المصاحبة إلا أن عدداً من الأبحاث المعاصرة (جونسون ١٩٨٢a، ١٩٨٢b) اقترح بأن التراجع عن البحث في الذكاء في ذلك الوقت فعلاً كان متسرعاً. فمثلاً؛ تعد علاقت الربط التي تقدر بحوالي ٤٠، والتي تربط كلاً من وقت رد الفعل المعقد والدرجات في امتحانات الذكاء علاقة ممكنة.

من جهة أخرى، لم يمض وقت طويل قبل ولادة اتجاه عملي لتقييم الذكاء، ففي عام ١٩٠٥ وضع الأخصائي النفسي الفرنسي ألفرد بينيه ومساعده ثيوفيل سيمون اختباراً للذكاء مبني على المهام التي يطلب من الأطفال في سن المدرسة أداؤها. ويطلق على هذا النوع من الاختبارات اليوم اسم "نسخة عن العمل" Job replica يتضمن الاختبار عينة من المهام التي يطلب عادة من الممتحن تأديتها. إن المهام الثلاثين التي يتضمنها مقياس بينيه – سيمون لاختبار الذكاء هذا قد تم رصدها في الجدول رقم ١–٢.

الجدول ١-٢

الاختبارات الثلاثون في مقياس بينيه - سايمون الأول للذكاء

١ . التناسق البصري

٢ . الإدراك الحسي

٣ . الإدراك البصري

٤ . إدراك الطعام

٥ . البحث عن الطعام ضمن ظروف غير موائمة

٦ . تنفيذ الأوامر البسيطة

٧ . التعرف على الأشياء بأسمائها

٨ . التعرف على الأشياء المصورة بأسمائها

٩ . تسمية الأشياء المصورة

١٠ . مقارنة خطوط الأشياء المختلفة على مستوى فوق عتبة الشعور

١١ . الذاكرة السمعية لأعداد مؤلفة من ثلاثة أرقام

١٢ . مقارنة أوزان الأشياء المختلفة على مستوى فوق عتبة الشعور

١٣ . التأثر بالآخرين

١٤ . تعريف الأشياء المألوفة

١٥ . تذكر الجمل

١٦ . تحديد الإختلافات بين الأشياء المألوفة الموجودة في الذاكرة

١٧ . تذكر صور الأشياء

١٨ . الرسم من الذاكرة

١٩ . الذاكرة السمعية لأعداد مؤلفة من أكثر من ثلاثة أرقام

٢٠ . تحديد التشابهات بين الأشياء المألوفة الموجودة في الذاكرة

٢١ . تمييز الخطوط

٢٢ . ترتيب خمسة أوزان

٢٣ . تحري الوزن المفقود

٢٤ . الأنغام

٢٥ . الكلمات المفقودة

٢٦ . تركيب الجمل

٢٧ . الإجابة على الأسئلة الإشكالية

٢٨ . حركة عقارب الساعة المتبادلة

٢٩ . الرسم من تصميم تم قصه من ورقة مطوية أربع طيات

٣٠ . التمييز بين المفاهيم المجردة

تجسد المهام المذكورة أعلاه فرضية "بينيه" بأن الذكاء يتألف من إمكانيات إدراكية متعددة تتغير مع تقدم الإنسان في العمر ويؤكد اختبار بينيه أيضاً على إيمانه بأن اختبارات الذكاء يجب أن تؤكد على القدرات الذهنية العليا. وعلى الرغم من أن معظم إختبارات بينيه تتألف من مهام مناسبة لطلاب المدارس (تنفيذ الأوامر البسيطة، المعرفة اللغوية، التعاريف، التعرف على الصور وإتمام الجمل) إلا أن بينيه استبعد كافة مقاييس المعرفة الإكاديمية من اختباره وركز على مقاييس الذكاء حصراً وذلك لأنه لم يربط الذكاء بأي علم يكتسب عن طريق الدراسة بل ربطه بالقدرة على الحكم والإستدلال والاستيعاب.

طبق بينيه وسيمون اختبار ذكائهم على عينة من الأطفال غير المتخلفين عقلياً أعمارهم ٣ و٥ و٧ و٩ و١١ عاماً بالإضافة إلى عينة من الأطفال والبالغين المتخلفين عقلياً. تم بعد ذلك ترتيب المهام تبعاً للعمر الذي يفترض فيه أن ينجز الطفل المهمة (مقياس العتبة العمرية)، ثم تم إجراء الاختبارات بدءاً بالمهمة ذات المقياس العمري الأدنى (المهمة ١) وانتهاءاً بالمهمة ذات المقياس العمري الأعلى (المهمة ٣٠). واستند بينيه وسيمون إلى النظام الفرنسي لتصنيف المتخلفين عقلياً

والذي يضم ثلاث فئات: الأغبياء والبلهاء والحمقى. وحدد بينيه وسيمون المهمة ٦ كالحد الأعلى للبالغين الأغبياء والمهمة ١٦ كالحد الأعلى للبالغين البلهاء والمهمة ٢٣ كالحد الإعلى للبالغين الحمقى.

نشرت نسخ مراجعة عن مقياس بينيه - سيمون عامي ١٩٠٨ و١٩١١ وحدثت تطورات أخرى على اختبارات الذكاء في بريطانيا والولايات المتحدة. وفي نسختي ١٩٠٨ و ١٩١١ من اختبار بينيه - سيمون تم تجميع المهام وفقاً للعمر المناسب لأدائها كما تمت إضافة مهام فرعية في حال نجاح نسبة من الأطفال تتراوح بين ٦٧٪ و ٧٥٪ في تجاوز هذه المهمة. هذه النسبة هي لأطفال من شريحة عمرية واحدة مع الأخذ بعين الاعتبار وجود نسبة قليلة من الأطفال الأصغر ونسبةعالية من الأطفال الأكبر يمكنها اجتياز هذه المهمة بنجاح. كان الاختبار يبدأ بتحديد العمر القاعدي Basal age للممتحن (وهو أعلى مستوى عمري للمهام الفرعية تمكن من تجاوزه بنجاح) ومن ثم تحديد العمر السقفي Ceiling age (وهو أدنى عمر فشل الممتحن في كافة مهامه الفرعية). ويحدد عمر الممتحن العقلي بربط كل مهمة فرعية ينجح فيها بعدد من الأشهر بين العمر القاعدي والعمر السقفي ثم يضاف الناتج إلى العمر القاعدي. وأخيراً، كانت تتم مقارنة العمر الناتج عن الاختبار بعمر الممتحن الحقيقي لتحديد ما إذا كانت قدراته العقلية أعلى أو أدنى من المعدل.

يعد مقياس ستانفورد - بينيه الذي نشره لويس تيرمان عام ١٩١٦ من أهم التعديلات على مقياس سيمون - بينيه والذي لم يكتف فيه تيرمان بمقارنة العمر العقلي بالعمر الفعلي للممتحن بل أدخل مبدأ الذكاء النسبي (ratio IQ) والذي يمكن تعريفه بنسبة العمر العقلي (بالأشهر) إلى العمر الحقيقي (بالأشهر) مضروباً بـ ١٠٠. وعليه فإن الأطفال ذوي معدلات الذكاء الطبيعية كانت معدلات ذكائهم ١٠٠، بينما تعدت معدلات ذكاء الأطفال الموهوبين الرقم ١٠٠ بشكل

ملحوظ وتدنت معدلات ذكاء الأطفال المتخلفين عقلياً دون ١٠٠ بشكل ملحوظ أيضاً.

أطلق اسم "الخبز والسمن لعلم النفس" على اختبارات الذكاء لأنها أتاحت الفرصة للعديد من علماء النفس لكسب المال بعيداً عن المختبرات والمكتبات ومدرجات الجامعات. فكثيراً ما كان علماء النفس هؤلاء يُرون وهم يتنقلون من مدرسة إلى أخرى ليختبروا ذكاء الأطفال قبل ضمهم إلى فصول تعليمية خاصة، كما كانوا يُشاهدون في عيادات إرشاد الأطفال حيث كانوا يحكمون ما إذا كان لمعدل ذكاء هؤلاء الأطفال المتدني دور في مشاكلهم السلوكية.

إن اختبار ستانفورد - بينيه اختباراً فردياً والذي خلافاً لاختبار المجموعات يعطى لكل فرد على حدة في وقت واحد، تم وخلال أقل من سنة من تاريخ نشر الاختبار تطوير اختبارين جماعيين للذكاء هما اختبارا الجيش ألفا وبيتا(٢)، وقد كان التركيز الرئيسي لاختبار ألفا هو على القدرة اللغوية إذ كان يتألف من: اختبار القدرة على اتباع الأوامر الشفهية واختبارالإستدلال الحسابي وإتخاذ القرارات العملية ومزاوجة الأضداد والمترادفات وترتيب الجمل المبعثرة وإكمال متسلسلات الأرقام والمتشابهات والمعلومات. وتجدر الإشارة هنا إلى أن الكثير من اختبارات القدرة الأكاديمية اليوم تحتوي على عناصر مشابهة. وعلى عكس اختبار ألفا، كان اختبار بيتا اختباراً لالغوياً يعطى للجنود الأميين أو غير الناطقين باللغة الإنجليزية وكان يتألف من عناصر بصرية - إدراكية حسية ومهارات حركية وذلك كمن يتتبع طريقاً الخروج في متاهة أو تصور عدد المباني

(٢) تم تطوير أول اختبار جماعي للذكاء من قبل بايل (١٩١٣). وكان اختبار بايل مؤلفاً من مقاييس للذاكرة المنطقية وذاكرة الحفظ عن طريق التكرار والقدرة على التعويض و بناء الكلمات وتداعي الأفكار وكان مرفقاً مع الاختبار معدلات للأداء تم الحصول عليها من عينات من طلاب المدارس. راجع بينتنر الاختبار عام ١٩١٧ وأصبح اسمه اختبار بينتنر- باترسون لمقاييس الأداء.

في رسم ثلاثي الأبعاد. وقد خضع أكثر من مليوني مجند لاختباري ألفا وبيتا بغرض تصنيفهم وتوزيعهم على المهام العسكرية المختلفة خلال الحرب العالمية الأولى وبعدها مباشرة (يركيس، ١٩١٩). وفي العشرينيات، خضع الملايين من طلاب المدارس في الولايات المتحدة لاختبار الذكاء الوطني وغيره من الاختبارات الجماعية المقتبسة عن اختباري الجيش ألفا وبيتا، كما استخدم اختبارانهما الاختبار العام لتصنيف الجيش واختبار التأهيل للقوات المسلحة في سنواتٍ لاحقة ويعد هذان الاختباران منحدرين من سلالة الاختبارين ألفا وبيتا.

لم يتم استخدام البيانات الناتجة عن اختباري ألفا وبيتا في رفع كفاءة العمليات الحربية لأسباب متعددة من أهمها انعدام الثقة لدى السلطات العسكرية في الأبحاث العلمية، إلا أن نتائج هذه الاختبارات وغيرها التي أجريت على مجموعات من المهاجرين إلى الولايات المتحدة تم تفسيرها بشكلٍ خاطئ لإثبات فكرة الدونية الفكرية لمواطني أوروبا الشرقية بالمقارنة مع المهاجرين من أوروبا الغربية والشمالية (بريغهام، ١٩٢٣؛ يركيس، ١٩٢١). وقد أدت أدلة تم الحصول عليها لاحقاً إلى تراجع بعض علماء النفس (مثل بريغهام، ١٩٣٠) عن تصريحاتهم حول الجنسيات وفروقات الذكاء. وبشكلٍ عام فإن نتائج الدراسات الأولى واللاحقة أشارت إلى تباين نتائج الأفراد في اختبارات الذكاء الأمريكية تبعاً لتشابه حضارتهم الأم مع الحضارة الأمريكية المسيطرة (أو إختلافها عنها).

بالإضافة إلى الاختبارات ذات التركيز اللغوي مثل اختبار ستانفورد – بينيه وغيره من اختبارات الذكاء الجماعية، تم في العقدين الأولين من القرن العشرين اعداد الكثير من الاختبارات الفردية ومقاييس الأداء وبطاريات الاختباربهدف تحديد مستوى القدرات العقلية للأفراد. ولا تزال العديد من هذه الاختبارات متوفرة حتى اليوم مثل لوائح سيغوين ومكعبات نوكس وتصاميم كوهس المكعبة

ومتاهات بورتيوس. وتتطلب اختبارات الأداء هذه، على عكس غيرها من الاختبارات التي تعتمد على القدرات اللغوية الشفهية والكتابية، التعامل مع الأجسام الفيزيائية كالمكعبات والمتاهات والخرز والصور المفككة (puzzles).

الاختبارات التحصيلية

لم تكن أدوات قياس الذكاء هي الاختبارات الوحيدة التي تم نشرها في بداية القرن العشرين، ففي عام ١٩١٠ تم تداول نماذج لاختبارات قياسية في الحساب والخط واللغة والتهجئة وغيرها من المواد الدراسية المدرسية. وفي عام ١٩٢٣ نشر ت. ل. كيلي الطبعة الأولى من اختبار ستانفورد للتحصيل، وفي عام ١٩٢٤ نشر إ. ف. ليندكويست امتحان آيوا للمحتوى الخاص بالمدارس الثانوية بينما نشر مجلس التعليم العالي (College Board) أول طبعة من اختبار الاستعداد المدرسي (SAT) ذو أسئلة الاختبارات المتعددة. وقد ساهم نمط الاختبارات القياسية ذات أسئلة الاختبارات المتعددة في تطوير العديد من اختبارات التحصيل القياسية خلال العشرينيات والثلاثينيات من القرن الماضي، كما ساعد استخدام الآلة في تصحيح الاختبارات في الثلاثينيات في تطور هذا النوع من الاختبارات بشكل كبير. ومما زاد من أهمية هذه الاختبارات تجميع معايير الاختبارات التي أجريت على عينات مختلفة تمثل المجتمع من طلاب المدارس ثم إجراء أبحاث للتأكد من صدق وثبات هذه الاختبارات بالإضافة إلى تحليل فقراتها المختلفة.

اختبارات الاستعداد

شاعت في العشرينيات والثلاثينيات، بالإضافة إلى اختبارات الذكاء والتحصيل، اختبارات تقيس الاستعداد الفردي في مجالات التوظيف وانتقاء الطلاب المتقدمين إلى المدارس والكليات المختلفة. وعلى عكس اختبارات الذكاء التي كانت مرتبطة بالمدارس فقط، تم تطوير اختبارات الاستعداد بغرض اختيار

وتصنيف الموظفين في مجالات الأعمال والصناعة والجيش. كما اثبتت هذه الاختبارات جدواها في المجالات المهنية، ففي كتابه الرائد: كيف تختار مهنتك، اقترح فرانك بارسون نموذجاً لمطابقة الأشخاص مع العمل المناسب لهم بناءً على مواصفات هؤلاء الأشخاص ومتطلبات الوظيفة. وقد كان بارسون على قناعة أن اختيار المهن وتوزيع القوة العاملة عليها يمكن أن يصبح أكثر فاعلية إذا ما أتيحت الفرصة للأفراد لمعرفة إمكاناتهم وقابلياتهم واهتماماتهم بالإضافة إلى معرفة متطلبات المهنة أو الوظيفة. وقد تلقت حركة بارسون للإرشاد المهني التي انطلقت عام ١٩٠٩ دعماً كبيراً من التطورات التي طرأت على اختبارات الاستعداد واختبارات الميول واختبارات الشخصية في العشرينيات والثلاثينيات. كما لعب نشر المعجم الأول للعناوين المهنية (DOT) عام ١٩٣٩ دوراً هاماً في دعم حركة بارسون أيضاً إذ يعتبر هذا المعجم الذي كان يعدل دورياً منذ تاريخ نشره أول محاولة حقيقية لتوفير نظام شامل لتصنيف المهن وتفصيل مقتضيات الوظائف المختلفة.

تضمنت اختبارات الاستعداد الأولى تلك التي عملت على قياس القدرات النفسية – الحركية والقابلية الميكانيكية والكتابية (السكريتارية) إلا أن اختبارات الاستعداد الفني والموسيقي لم يتم إعدادها حتى الثلاثينيات. وكان كارل سيشور قد نشر مقياس الموهبة الموسيقية عام ١٩١٨ في حين طور جونسون أوكونور في العشرينيات مقياساً للمهارة الحركية وطور د. ج. باترسون وزملاؤه في جامعة مينيسوتا في أواخر العشرينيات وبداية الثلاثينيات اختبارات للقدرة الميكانيكية والمكتبية (السكريتارية). وقد استخدمت جميع هذه الاختبارات وغيرها في تقديم الإرشاد المهني وانتقاء الموظفين لعقدين كاملين.

بالإضافة إلى اختبارات الاستعداد الواحد، تم نشر بطارية اختبارات الاستعداد العام واختبارات الاستعداد التبايني وهي اختبارات لا تقيس

استعداداً معيناً فحسب بل عدداً من الاستعدادات في آنٍ واحد. وقد اثبتت هذه الاختبارات جدواها في اختيار الموظفين للمهام اللاإدارية كالكتاب وعمال خطوط التجميع في المصانع وعمال الصيانة. كما اثبتت اختبارات الاستعداد تفوقها على أدوات تقييم الأداء أثناء العمل في توقع نجاح الأفراد في البرامج التدريبية المختلفة. ومن أفضل المقاييس الخاصة بالاستعداد تلك التي تقيس السرعة والدقة والتي تستخدم في اختيار موظفين للأعمال المكتبية وفي تقييم الإدراك الميكانيكي والعلاقات الفراغية عند اختيارعمال خطوط الإنتاج وعمال الصيانة. كما ساهمت اختبارات الذكاء العامة في اختيار العاملين في مجالات مختلفة إذ أكد فيليب فيرنون (١٩٦٠) على أن الذكاء العام لا القابليات أو الاستعدادات الخاصة هو الذي يحدد النجاح المهني.

أحداث تاريخية أخرى

ومن الجدير بالذكر وجود أحداث أخرى هامة في تاريخ الاختبار النفسي والتعليمي كنشر مراجعات لمقياس ستانفورد – بينيه للذكاء عامي ١٩٣٧ و١٩٦٠ ونشر مقاييس ويكسلر للذكاء (للأطفال والبالغين والأطفال دون سن المدرسة) ونشر مقاييس بايلي لتطور الرضع. ومن الأحداث الهامةفي تاريخ اختبارات التحصيل ظهور العديد من الطبعات المعدلة عن الاختبارات الموجودة وظهور اختبارات جديدة وتطويرعمليات التقنين ونمو مؤسستي "خدمة الاختبارات التعليمية" و"برنامج الاختبار للتعليم العالي في الولايات المتحدة" وغيرها من المؤسسات المهنية للاختبارات. ومن أهم نتائج البحث في هذا المجال التوصل إلى حقيقة أن اختبارات الاستعداد غير منحازة ضد الأقليات ولها درجة عالية من الصدق في توقع النجاح المهني للأفراد في مواقع متعددة (شميدت وهنتر، ١٩٧٧، ١٩٧٨، ١٩٨١).

ويجب التأكيد كذلك على التقدم في مجالي نظرية وطرق الاختبار النفسي والتعليمي، فمع نهاية العشرينيات كان تشارلز سبيرمان و ترومان كيللي وغيرهما من علماء النفس قد وضعوا أساسيات نظرية الاختبار الكلاسيكية. وقد تزامنت هذه التطورات مع مساهمات ومساعدات إحصائية في مجالي النظرية وطرق البحث قدمها كارل بيرسون ور. أ. فيشر وغيرهما. لقد كان للتطورات في تحليل العوامل وتحليل البنى الكامنة وتحليل المسارات ونظرية الإستجابة للفقرة وغيرها من التطورات والاكتشافات في مجال الاختبارات دور كبير في تطوير وتوسيع نظرية الاختبار النفسي وطرق البحث فيها خلال القرن العشرين. وبدورها فقد ساهمت ثورة التكنولوجيا وتقنية الكمبيوتر ليس في تحليل الامتحانات وتدوين نتائجها فحسب بل في إجراء هذه الامتحانات وتأويل معطياتها مما أخذ مساهمات الاختبارات النفسية في مجالي العلوم والتطبيق إلى أبعاد لم يتخيلها أكثر المؤيدين لهذه الاختبارات. وقد كان للحاسب أثر مشابه لأثر الإحصاء على عالم الاختبارات في النصف الأول من القرن العشرين.

تصنيف الاختبارات ومصادر المعلومات

كما أشرنا سابقاً فإن أدوات التقييم النفسي يمكن أن تصنف تبعاً إلى ما تهدف هذه الأدوات إلى قياسه كالقدرات الإدراكية أو العاطفية اللاإدراكية أو النفسية – الحركية. ويمكن تقسيم كلٍ من هذه الفئات إلى فئات فرعية أخرى إذ يمكن للاختبارات الإدراكية أن تقيس الذكاء والتحصيل والقدرات الخاصة بينما تقيس اختبارات الشخصية الطباع والصفات الشخصية وغيرها من الخصائص العاطفية. كما يمكن وصف اختبارات الذكاء والاستعدادات الخاصة بأنها اختبارات لطرق تفكير تجميعية وتفكيكية (إبداعية) أو اختبارات قدرة أكاديمية (مدرسية) أو اختبارات غير منحازة لحضارةٍ ما على حساب أخرى. ويمكن

تقسيم هذه الاختبارات بناءً على الأهداف التي تخدمها إلى اختبارات اختيار للمتقدمين وتحديد الأدوار(اختبارات التنسيب) واختبارات لتصنيف وتقييم الأداء واختبارات للجاهزية وأدوات للتحليل والتنبؤ. ويمكن لهذه الاختبارات أن تكون قياسية مما يقتضي استخدام معاييـر Norms ثابتة لتحليل النتائج أو غير قياسية (يضعها المدرسون). وهنالك أيضاً اختبارات للسلوك والإرتقاء واختبارات عصبية – نفسية واختبارات حركية – حسية واختبارات للسمع والنطق واختبارات تكييفية واختبارات المتشابهات واختبارات الفقرات الأساسية المتكررة واختبارات تسلسلية وغيرها. ويمكن تصنيف الاختبارات وفقاً للتوقيت الزمني الذي يتم فيه الاختبار ضمن العملية التعليمية أو التدريبية أو الشفائية أو العلاجية ويطلق على عمليات التقييم التي تفيد في تقييم التطور الحاصل أثناء تطبيق برنامجٍ ما عمليات التقدير التكويني (Formative Assessment) بينما يطلق على العمليات التي تقييم النواتج النهائية عمليات التقدير التراكمي (Summative Assessment).

طرق أخرى لتصنيف الاختبارات

يمكن تصنيف الاختبارات تبعاً لطرق وضع هذه الاختبارات والشكل الذي تتخذه وطرق إجرائها وطرق الإجابة على أسئلتها وطرق وضع درجاتها وطرق تحليل الدرجات أو تقييم الأداء.

تتنوع طرق وضع الاختبارات بتنوع الأهداف المراد تحقيقها منها ولكنها غالباً ما تنطلق من مخطط مفصل أو جدول للمواصفات يحدد محتوى الاختبار، ثم يتم إتخاذ قرارات متعلقة بطبيعة الأسئلة التي سيتضمنها الاختبار: أسئلة موضوعية (اختبارات متعددة أو أسئلة الخطأ والصواب أوأسئلة الإجابات القصيرة أو أسئلة المطابقة) أو أسئلة ذات إجابات مفتوحة (مقالية) أو أسئلة

تطبيق عملية. ويمكن ترتيب هذه الأسئلة ضمن عدد من الاختبارات الفرعية أو على شكل كتيب لولبي (أي أسئلة مرتبة بالتناوب حسب نوعها ومن الأسهل إلى الأصعب). وتتطلب الإجابة على أسئلة الاختبارات المتعددة قدرات معرفية واستيعابية وتطبيقية وتحليلية وتجميعية وتقييمية وغيرها من العمليات الإدراكية (بلووم وكاثول، ١٩٥٦).

ويمكن لأسئلة الاختبارات أن تكون بصرية (مطبوعة ضمن كتيب أو على بطاقات أو على شاشة الكمبيوتر) أو سمعية (يقرؤها المراقب الموجود في قاعة الاختبار أو تقدم على شريط مسجل). كما تطبع تعليمات الاختبار ضمن كتيب الاختبار أو على شاشة الحاسب أو تقرأ من قبل المراقب الموجود في القاعة أو يشار إليها بالإيماء. وقد تكون أسئلة الاختبارات في حد ذاتها أسئلة سهلة ولكنها عديدة لا يتمكن معظم الممتحنين من إكمالها ضمن الوقت المحدد للاختبار (اختبار سرعة) أو يكون الوقت المحدد للاختبار طويلاً جداً إلا أن الأسئلة صعبة(اختبار قوة).

يضع الممتَحَنون (ووفقاً لشكل الاختبار وأغراضه) إشارة أمام الإجابة الصحيحة على ورقة الأسئلة أو على ورقة إجابة منفصلة أويطبعون الإجابات مستخدمين لوحة مفاتيح الكمبيوتر أو يجيبون على الأسئلة شفهياً أو بالإشارة أو بالتطبيق أو بغيرها من الطرق التي تسمح لهم بالإجابة. ويمكن للإجابات أن تكون محددة بعدة خيارات (أسئلة الاختيار من المتعدد) أو مفتوحة كالأسئلة المقالية أو أسئلة المشاريع التطبيقية.

وتعتمد عملية منح الدرجات في الاختبارات المختلفة على كون الاختبار موضوعياً أو ذاتياً ففي الاختبارات الموضوعية هنالك عدد محدد من الإجابات ومفتاح للأجابات يمكن مقارنة جميع الإجابات به. وفي حال عدم وجود أي أخطاء مكتبية تبقى درجات الممتحن ثابتة حتى مع اختلاف واضعي الدرجات. أما في الاختبارات ذات الطابع الذاتي كتلك التي تتضمن أسئلة مقالية، تختلف

الدرجات المعطاة للممتَحَن باختلاف المصحح[٣] على الرغم من وجود تعليمات وإرشادات عامة تنظم عملية وضع الدرجات. وينطبق هذا الكلام على تقييم المشاريع والعروض التجريبية وغيرها. وبغض النظر عن أشكال الاختبارات أو أسئلتها أو عن كون الاختبارات موضوعية أو ذاتية، غالباً ما تعتمد درجات الاختبارات على معادلة ١ - صفر (أي تمنح درجة صفر للإجابة الخاطئة وعدد من الدرجات وغالباً ١ للإجابات الصحيحة)، كما يمكن أن توزع الدرجات على الأسئلة وفقاً لصعوبة أو أهمية السؤال. وفي بعض الاختبارات المقالية والمشاريع والتجارب العملية تعطى درجات للدقة ومحتوى الإجابة ولشكل الإجابة أيضاً (القواعد، الخط، الكفاءة، سلاسة الأسلوب... إلخ).

ومن الجدير بالذكر أن عملية وضع الدرجات ليست هي نهاية المطاف في عملية الاختبار إذ يجب تقييم العناصر التي منحت درجات والعناصر التي لم تمنح أية درجات في كل اختبار وتحليلها وفق معايير معينة (ذاتية أو موضوعية). وفي معظم الاختبارات المدرسية يقوم المدرس بوضع الدرجات وتقييم نتائج الاختبارات وتحليل معطياتها دون الإستعانة بأي شخصٍ آخر. فالمدرس إذاً هو واضع الدرجات وهو الذي يحدد التقدير النهائي للطالب تبعاً لمقاييس تجربته الخاصة، وقد يكتب هذا المدرس تعليقاً مفيداً على أداء الطالب في ذلك الاختبار أو لا يكتب. أما في الاختبارات التحصيلية الرسمية، يفرق المختصون بين الاختبارات التي تمنح الدرجات بناءً على مرجعية معيارية (Norm-referenced Assessment) والاختبارات التي تمنح الدرجات بناءً على مرجعية المحك (Criterion-referenced Assessment). ونظراً لأن الاختبارات

---

(٣) استخدمت كلمة مصحح هنا على الرغم من أن واضع الدرجات هي المصطلح الأنسب إذ أن الذي يضع الدرجات لا يصحح الإجابات بل يقييمها فقط. (المترجم).

معيارية المرجع يتم قياسها من خلال إجراء الاختبار على عينة من الأشخاص تمثل المجتمع، يتم تفسير (تحليل) درجات هذه الاختبارات بمقارنتها مع المعايير الناتجة عن عينة القياس. أما درجات الاختبارات المحكية المرجع (والتي تسمى أيضاً اختبارات الإتقان Masterytests)، فيتم تفسيرها عن طريق مقارنتها بأهداف أو مستويات محددة مسبقاً للأداء المقبول. ويقوم بتفسير نتائج الاختبارات عادةً أخصائيون في مادة الاختبار. كما يتم تفسير نتائج العديد من الاختبارات النفسية واختبارات القدرة والشخصية بواسطة الحاسب، ويحتوي برنامج تحليل الاختبارات بواسطة الحاسب (CBTI) على عدد من الجمل التي تصف إجابات معينة على الأسئلة أو الدرجات النهائية للاختبار بكامله، ومن ثم تتم طباعة هذه الجمل في وثيقة توضح أداء الفرد في الاختبار.

مصادر للمعلومات عن الاختبارات

لا يوجد مصدر واحد يحتوي على قائمة ووصف لكل الاختبارات النفسية والتعليمية، إلا أن أشمل قوائم الاختبارات المطبوعة موجودة في كتاب الاختبارات المطبوعة ٤ (مورفي، كونولي، وإيمارا، ١٩٩١) وكتاب اختبارات (سويتلاند وكيسير، ١٩٩١). ويحتوي برنامج F-8 على قوائم خاصة بالاختبارات التحصيلية بينما يحتوي برنامج H-7 على قوائم باختبارات الاستعدادات العامة والخاصة. وتعد الكتب الثلاثة التالية من أهم المراجع الخاصة بالاختبارات غير المنشورة: دليل المقاييس العقلية غير المنشورة (غولدمان وميتشل ١٩٩٠، غولدمان وأوزبورن ١٩٨٥)، دليل المستهلك للاختبارات المنشورة (هاميل، براون وبراينت، ١٩٩٢)، وفهرس الاختبارات المستخدمة في أطروحات الدراسات العليا (فابينو، ١٩٨٩). كما يمكن الإستفادة من قواعد البيانات الالكترونية مثل خدمة تجميع

الاختبارات التابعة لـ (ETS مجموعة الاختبارات وخدمة الاختبارات التعليمية في برينستون نيوجرسي) وقاعدة بيانات PsycINFO.

ولقراءة مقالات نقدية عن الاختبارات المختلفة يمكن العودة إلى كتب المقاييس العقلية السنوية بمجلداتها المختلفة (كونولي وإمبارا، ١٩٩٥، كونولي وكرامر، ١٩٨٩، وكرامر وكونولي، ١٩٩٢) وكتاب مقالات نقدية عن الاختبارات (كيسير وسويتلاند، ١٩٨٤-١٩٩٤) وكتاب الاختبارات في مجال التعليم: آراء النقاد والمقيمين (ليفي وغولدشتاين، ١٩٨٤) وكتاب اختبار الأطفال (ويفر، ١٩٨٤) وكتاب اختبار المراهقين (هارينغتون، ١٩٨٦) وكتاب اختبار البالغين (سفايرسنسكي، ١٩٨٥). وتحتوي الكتب التالية على قوائم ووصف لمقالات المجلات التخصصية المتعلقة بالاختبارات وغيرها من الفصول ضمن الكتب بشكل عام والكتب المعنية بالاختبارات بشكل خاص: ملخصات أطروحات الدراسات العليا في مجال علم النفس، فهرس التعليم، الفهرس الحالي للمجلات التخصصية التعليمية. كما تحتوي بعض المجلات التخصصية في علم النفس والتعليم والأعمال وغيرها على مقالات تتعلق بالاختبارات وعمليات الاختبار. وتنشر المجلات التخصصية التالية مقالات عن المهارات الإدراكية والحركية - النفسية: المجلة الأمريكية للأبحاث التعليمية ومجلة المقاييس النفسية التطبيقية ومجلة المقاييس النفسية والتعليمية ومجلة علم النفس السريري ومجلة علم النفس الاستشاري ومجلة المقاييس التعليمية ومجلة علم النفس المدرسي ومجلة السلوك المهني ومجلة القياس والتقييم في الإرشاد والتطوير ومجلة علم نفس الموارد البشرية ومجلة التقييم النفسي - التعليمي ومجلة التقييم النفسي ومجلة علم النفس في المدارس.

وتتوفر قوائم لبرامج الحاسب التي تستخدم في مجال الاختبار في كتاب استخدام الحاسب في علم النفس: دليل البرامج (ستولوف وكاوتش، ١٩٩٢)

وكتاب مصادر برامج الكمبيوتر النفسية (كروغ، ١٩٩٣)، ويحتوي الملحق (ج) من هذا الكتاب على قائمة بعناوين الشركات التي تسوق برامج الحاسب المستخدمة في عمليات الاختبار التعليمي والنفسي.

كما توفر القوائم التي تنشرها الشركات المتخصصة في إنتاج مواد الاختبارات مصدراً وإن كان محدوداً للمعلومات المتعلقة ببعض الاختبارات الخاصة. ويحتوي الملحق (ب) من هذا الكتاب على قائمة بعناوين وأرقام هواتف دور نشر وتوزيع مواد التقييم التعليمية والنفسية، وتتوفر معلومات تفصيلية عن معظم الاختبارات المنشورة في أدلة المستخدم التي ترفقها دور النشر عادةًبالاختبارات وأدوات القياس المختلفة.

المقاييس والاعتبارات الأخلاقية والقانونية

ساهمت عدة عوامل خلال العقود القليلة الماضية في زيادة أهمية أدوات القياس النفسية والتعليمية وفي توطيد الأسس العلمية لعملية الاختبار، ولعل أهم هذه العوامل هو تناول موضوع القياس النفسي والتعليمي في الآلاف من الكتب والمقالات والفصول ضمن الكتب والمئات من المحاضرات التي ألقيت في مؤتمرات إقليمية ودولية. وقد ساهم التأكيد على أخلاقيات ومعايير الاختبار من قبل هيئات رسمية كالإتحاد الأمريكي لعلم النفس والإتحاد الأمريكي للبحوث التعليمية والإتحاد الأمريكي للإرشاد وشؤون الموظفين في ضمان تطور ونمو مجال الاختبارات التعليمية والنفسية بطريقة أكثر انتظاماً وأقل فوضويةً من السابق. كما يجري حالياً تطويرعدد من الإجراءات وأدوات القياس السايكومترية التي لا تستهدف الربح المادي فحسب بل تحقق درجات عالية من الصدق والثبات ويتم تطبيقها بشكلٍ يضمن مصالح الناس.

تُجرى الاختبارات عادةً لعدة أغراض في المدارس والجامعات والعيادات والمحاكم ومجالات الأعمال والصناعة والدولة والجيش وغيرها من المجالات،

وتوفر معدلات الدرجات النهائية الناتجة عن هذه الاختبارات معلومات هامة تساعد في إتخاذ قرارات بشأن أشخاص وبرامج مختلفة (تعليمية أو غيرها). ويمكن تلخيص استخدامات نتائج الاختبارات النفسية والتعليمية في النقاط التالية:

- تصنيف وتوزيع واختيار الأفراد في المؤسسات التعليمية والمهنية.
- فَصْل أو ترقية أو الإبقاء على الطلاب في المدارس والموظفين في المؤسسات.
- منح الحرفيين أو أصحاب المهن تراخيص بمزاولة مهنهم.
- تشخيص المشكلات النفسية أو التعليمية والتعرف على احتياجات المصابين بغرض توفير المساعدة لهم عن طريق جلسات الإرشاد أوالعلاج أوغيرها من الإجراءات.
- تقييم فعالية التدريب أو العلاج أو غيرها.
- توفير أدوات لقياس متغييرات السيطرة والمتغيرات المستقلة وغير المستقلةللأبحاث في تأثير الإجراءات العلاجية على التغيرات السلوكية.
- توفير معلومات للأفراد تساعدهم على معرفة أنفسهم بشكلٍ أفضل.

وتعتمد درجة الوثوق في تحقيق أهداف الاختبارات المذكورة أعلاه على إجراء اختبارات صادقة وباستخدام إجراءات قياسية موحدة مع مراعاة تناسب الاختبارات مع الأفراد موضع الاختبار والأهداف المراد تحقيقها من الاختبار.

مقاييس الاختبار

يعد الدليلان التاليان: مقاييس للاختبارات النفسية والتعليمية (الإتحاد الأمريكي للبحوث التعليمية، ١٩٨٥) وتوجيهات للاختبارات المحوسبة وتفسير

نتائجها (الإتحاد الأمريكي لعلم النفس، ١٩٨٦) مصدرين هاميين لمطوري ومستخدمي ومسوقي الاختبارات. ويوفر دليل "المقاييس" معايير لتقييم الاختبارات والإجراءات الاختبارية وآثار استخدام الاختبارات. كما يحتوي هذا الدليل على الأهداف المعيارية الأولية التي يجب على واضعي الاختبارات والقائمين على الاختبارات تحقيقها، بالإضافة إلى الأهداف المعيارية الثانوية التي يُشجع على تحقيقها مع العلم أن هذا قد لا يكون ممكناً في حالاتٍ معينة. وتقسم الأهداف المعيارية إلى أربع مجموعات تتعلق ببنية الاختبار وطريقة تقييمه وبالمعايير المهنية لاستخدام الاختبارات (في المجالات السريرية والتعليمية والإرشادية والتوظيفية ومجالات منح الشهادات والتراخيص وتقييم البرامج) وبالتطبيقات الخاصة (كاختبار الأقليات اللغوية والأشخاص ذوي الإعاقات) وبالإجراءات الاختبارية.

وبالإضافة إلى المبادئ المتعلقة بواجبات مستخدمي ومطوري الاختبارات الموجودة في كتاب المقاييس، يوفر كتاب توصيات للاختبارات المحوسبة قائمة مؤلفة من تسع توصيات لإجراء الاختبارات عن طريق الحاسب وتفسير نتائجها، كما يوفر٣١ توصية لمطوري خدمات الاختبارات المحوسبة (مثل: العوامل البشرية والخصائص السايكومترية والتصنيف ومصداقية تفسير الدرجات عن طريق الحاسب). ولم تقتصر المعايير والتوجيهات على مساعدة مطوري ومستخدمي الاختبارات فحسب بل ساهمت في حث الهيئات المسوقة للاختبارات على التأكد من أن الأفراد الذين يشترون الاختبارات منهم مؤهلون تماماً لاستخدامها. ومن الجدير بالذكر أن دستور الممارسات الاختبارية العادلة في مجال التعليم (فيرمر ودايموند وكامارا، ١٩٨٩)، والذي بذل جهداً كبيراً للعمل مع ناشري الاختبارات، ساهم في تحسين نوعية الإجراءات الاختبارية بما يخدم المصلحة العامة إذ يحدد ويقسم الدستور مسؤوليات مطوري ومستخدمي الاختبارات إلى أربع

فئات: تطوير واختيار الاختبارات المناسبة وتحليل النتائج والسعي لتحقيق العدالة وإبقاء الممتحنين على دراية بكل ما يتعلق بالاختبار.

تختلف مؤهلات مستخدمي الاختبارات مع إختلاف أنواع الاختبارات وأغراضها فعلى سبيل المثال إن المؤهلات المطلوبة لشراء اختبار فردي أشد صرامة من تلك المطلوبة لشراء اختبار جماعي، كما أن المؤهلات المطلوبة لشراء اختبارات الذكاء أعلى من تلك المطلوبة لشراء الاختبارات التحصيلية واختبارات الاستعدادات الخاصة. وقد نهجت عدة منظمات نهج الشركة المتحدة لعلم النفس والتي اعتمدت نظاماً خاصاً يحدد ما يطلب من الراغبين بشراء الاختبارات، ويقسم هذا النظام الاختبارات إلى ثلاثة مستويات (أ، ب، ج) مرتبة تصاعدياً من حيث المتطلبات إذ يمكن للمدارس وغيرها من المنظمات أن تشتري اختبارات من الفئة (أ) بعد تقديم رقم شراء رسمي أو رسالة رسمية موقعة من قبل شخص مسؤول عن المشتريات ضمن المؤسسة. وأما بالنسبة للأفراد الراغبين في شراء اختبارات من الفئتين (ب) و (ج) فيطلب منهم إرسال إثبات عن مؤهلاتهم قبل أن يسمح لهم بشراء المواد الاختبارية المصنفة ضمن هاتين الفئتين، فيشترط في الراغبين بشراء اختبارات من الفئة (ب) حيازتهم على درجة الماجستير على الأقل في علم النفس أو في التعليم أو ما يعادلها من التدريب المتعلق بالتقييم النفسي أو أن يكونوا أعضاء في أي إتحاد مهني يخضع أعضاؤه لتدريب مناسب في مجال التقييم. أما الراغبون في شراء اختبارات من الفئة (ج) فيجب أن يكونوا من حملة درجة الدكتوراة في التعليم أو علم النفس أو ما يعادلها من التدريب في مجال التقييم أو أن يكونوا حاصلين على شهادة أو إجازة تتطلب تدريباً وخبرة مناسبين في مجال التقييم النفسي. ولقد وضعت عدد من دور النشر (كدار نشر علماء النفس الاستشاريين وNCS للتقييم ومصادر التقييم النفسي وغيرها من دور النشر المذكورة في الملحق ب من هذا الكتاب) نظاماً مشابهاً لمؤهلات

الراغبين بشراء الاختبارات النفسية. وعلى الراغبين في شراء الاختبارات ملء استمارة لتحديد درجات مؤهلاتهم.

المبادئ والقواعد الأخلاقية

وضع دستور الممارسات الاختبارية العادلة في مجال التعليم (فيرمر ودايموند وكامارا، ١٩٨٩) لائحة بعدد من الأخلاقيات المرتبطة بمطوري ومستخدمي أدوات الاختبار، بينما حددت وثيقة أخرى هي "المبادئ الأخلاقية لعلماء النفس ودستور سلوكهم" (الاتحاد الأمريكي لعلم النفس، ١٩٩٢) سلوك علماء النفس كعلماء وكممارسي مهنة وكمحافظين على المصلحة العامة. وتؤكد هذه الوثيقة على أن التقييم والتشخيص الرسميين يجب أن يتما في ظروف رسمية فقط وعلى أيدي مستخدمين للاختبارات مدربين وأكفاء. وينص هذا الدستور الأخلاقي على استخدام طرق علمية في تصميم وإختيار الاختبارات المناسبة لمجموعة معينة من السكان. ثم يتم بعد ذلك إجراء الاختبارات ضمن ظروف آمنة ثم توضع الدرجات وتحلل النتائج بطريقة حكيمة.

ومن الجدير بالذكر أن توفر اختبارات مصممة بدقة وحذر ووجود معايير ثابتة لتوزيعها لا يضمن أن عمليتي إجراء الاختبارات النفسية والتعليمية وتحليل نتائجها تتمان بشكل صحيح. وعلى الرغم من أن علماء النفس يعترفون بأن مسؤولية استخدام الاختبارات بشكل صحيح تقع على عاتق مستخدمي الاختبارات أنفسهم، تنقص العديد من المرشدين الاجتماعيين وأصحاب العيادات وغيرهم من العاملين في مجال الخدمات البشرية، المعرفة والمهارات اللازمة لإجراء وتحليل بعض أدوات القياس السايكومترية. ويجب إعلام أولئك الذين لم يحصلوا على التدريب الكافي بالنقص المعرفي لديهم ويجب بالتالي تشجيعهم على الإلتحاق بدورات تدريبية أو العمل مع المختصين والخبراء والرجوع إلى مصادر المعلومات المتوفرة بغرض تدعيم إمكاناتهم. كما يجب على القائمين على

الاختبارات أن يكونوا على درجة من الحساسية والدراية بلمتطلبات وحاجات الممتحنين والمجتمع بشكل عام. ويعرض كتاب "الاستخدام المسؤول للاختبارات" دراسات لعدة حالات تم فيها إجراء عدد من الاختبارات ووضع درجاتها ثم تحليلها بشكل خاطئ. ومن الجدير بالذكر أن الأمثلة المطروحة في الكتاب الآنف الذكر لا تقتصر على مادة دراسية واحدة بل تشمل عدداً لا بأس به من العلوم والمقررات.

ومن أهم المبادئ التي يتوجب على القائمين على الاختبارات والممتَحَنين معرفتها مبدأ الموافقة المطلعة ومبدأ السرية المهنية. لطالما ارتبط مبدأ الموافقة المطلعة بعالم الطب فقط حيث يعطي المرضى موافقتهم على إجراء عملية جراحية ما لهم (أو غيرها من المعالجات) أو على المشاركة بدراسة سريرية ما بعد الإطلاع على جميع الحقائق الطبية والمخاطر المرتبطة بما هم مقدمون عليه. أما استخدام هذا المصطلح في مجال علم النفس وفي مجال الاختبارات النفسية تحديداً فيشير إلى موافقة الفرد على المشاركة في اختبار نفسي أو المشاركة في دراسة نفسية بعد الإطلاع على جميع الحقائق والمخاطر التي يمكن أن تترتب على هذه المشاركة. وقبل إجراء معظم الاختبارات النفسية الرسمية يتم الحصول على الموافقة المطلعة للممتحن أو ممثله القانوني. ويفترض هنا أن المعلومات الشخصية التي يتم الحصول عليها من هذه الاختبارات لا تتم مناقشتها إلا مع الممتحن نفسه فلا يطلع عليها أي طرف ثالث دون موافقة الممتحن. وعليه تصبح المعلومات "مصانة" أي محمية من أن يطلع عليها من لا يملك الحق القانوني بالإطلاع. ويوضح الصندوق ١-١ نموذجاً للإستمارة التي يملؤها الممتحن للموافقة المطلعة على المشاركة في اختبار نفسي. تفيد الاستمارة بأن الممتحن على دراية بالأهداف الأساسية للاختبار وأن بإمكانه الإنسحاب في أي وقت وأن نتائج الاختبار لن يطلع عليها غيره دون موافقته. إذن، تشير الاستمارة إلى أنه تم

إعلام الممتحن بهدف الاختبار وبمن يمكنه الحصول على نتائجه وكيف سيتم استخدام هذه النتائج.

يعود أصل مبدأ السرية المهنية إلى معلومات يأتمن المدعي أو المدعى عليه في قضية ما محاميه أو طبيبه أو رجل الدين أو الزوج (أو الزوجة) عليها، وتصبح هذه المعلومات محصنة من الإدلاء بها أمام القضاء إذا أراد الموكل أو المريض أو التائب أو الشريك في الزواج هذه الحصانة.

الصندوق ١-١

الموافقة المطلعة على اختبار نفسي

أنا الموقع أدناه ------------------ أعطي موافقتي الطوعية على أن أكون أحد المشاركين في الاختبار النفسي الذي تجريه ---------------. كما أقر بأنني قد أُعطيت تفسيراً كاملاً وواضحاً عن الطبيعة العامة للاختبار وأهدافه المحددة والأسباب التي سأختبر لأجلها، وأنني قد أُعلمت بأنواع الاختبارات وغيرها من الإجراءات التي سيتم إجراؤها والطريقة التي سيتم استخدام نتائج الاختبار فيها.

وأنا على علم أنه من غير الممكن لمجري الاختبار أن يوضح جميع أوجه الاختبار قبل إتمامه بكامله. كما أنني على معرفة بأنه يمكنني الإنسحاب من الاختبار في الوقت الذي أشاء دون أن أتعرض لأي خسارة أو جزاء. وأعلم كذلك أنه سوف يتم إعلامي بنتائج الاختبار عند توفرها ولن يطلع عليها أي طرف آخر دون موافقتي. كما أود طلب إرسال نسخة من النتائج إلى العنوان التالي:

اسم الممتَحَن ------------------ التوقيع --------------------

التاريخ ---------------- توقيع مجري الاختبار ----------------

ويستخدم مصطلح السرية المهنية في السياق النفسي بمعنى أن المعلومات الناتجة عن الاختبارات النفسية يمكن إطلاع أطراف ثالثة عليها بطلب رسمي فقط، وحتى في حال موافقة الممتحن على إطلاع أطراف ثالثة على نتائج الاختبار من الممكن أن تبقى المعلومات محصنة من الناحية القانونية بحيث لا يطلع أحد غير الممتحن، أو في حال كون الممتحن قاصراً يمكن لكل من ولي أمره أو وكيله القانوني أو طبيبه أو طبيبه النفسي، الاطلاع على المعلومات الناتجة عن الاختبار.

ويوجد العدد من القوانين (مثل: الحقوق التعليمية للعائلات وقانون الخصوصية لعام ١٩٧٤) التي تنص على سرية المعلومات الناتجة عن الاختبارات النفسية والتعليمية المختلفة، وأن نتائج الاختبارات وملفات الطلاب المحفوظة لدى المدارس والمؤسسات التعليمية التي تتلقى دعماً مادياً من الموارد الفيدرالية لا يمكن الإطلاع عليها إلا بموافقة الطالب أو وليه الخطية وبطلب شخصي من الجهة الراغبة في الإطلاع على هذه الوثائق. أما بالنسبة للوالدين والعاملين في المدرسة الذين لهم "اهتمام تعليمي مشروع" ببعض الطلاب، يسمح لهم هذا القانون والفقرات ٩٤-١٤٢ من القانون العام في حالة الطلاب المعاقين، بمراجعة ملفات هؤلاء الطلاب.

وبالإضافة إلى حَقَّيْ الموافقة المطلعة والسرية المهنية، يحق للأشخاص المختبرين نفسياً أن توصف حالتهم بشكل صحيح ولكن بمصطلحات تخلو من التشهير والإهانة إذ يتم اختيار مصطلح لوصف الحالة خالٍ من الدلالات السلبية، فعلى سبيل المثال إن استخدام مصطلحات مثل متخلف عقلياً أو ضعف عقلياً له وقع أقل إدانةً من مصطلحات غبي أو أبله أو أحمق، ولذلك يفضل استخدام المصطلحين الأوليين. كما أن للأشخاص المعاقين جسدياً الحق في أن يتم اختبارهم باختبارات سايكومترية مصممة بشكل خاص للشريحة السكانية التي ينتمون إليها والتي تم قياسها وتوثيقها استناداً لعينات من الأشخاص مختارة من تلك الشريحة السكانية. ويعد هذا أحد تعاريف العدالة أوالتحرر من التحيز في الاختبار.

القضايا القانونية في الاختبارات الوظيفية

ليس من المستغرب أن عدد القضايا القانونية المتعلقة بالاختبارات التعليمية والنفسية قد ارتفع بشكل كبير نتيجة الاستخدام الواسع النطاق لهذه الاختبارات في مجالات التوظيف والتعليم والعيادات النفسية ونتيجة لكون نتائج هذه الاختبارات ذات تأثير عميق على حياة الأفراد ونتيجة للتخوف على حقوق النساء والأقليات العرقية.

الجدول ١-٣

قضايا وأحكام قانونية هامة متعلقة بالاختبارات النفسية ضمن مجالي التوظيف والتعليم

١٩٦٣ **مايرت ضد موتورولا:** إن اختبار التوظيف رقم ١٠ الذي تستخدمه شركة موتورولا لا يدخل تحت مظلة تكافؤ الفرص ولا يأخذ بعين الاعتبار المجموعات البشرية المحرومة والفقيرة. كما أن هذا الاختبار قديم بالٍ لا يتماشى مع الوقت الحاضر وقد تم قياسه بناءً على معايير norms تم الحصول عليها عن

طريق إجراء الاختبار على عينة من الناس تنتمي إلى الفئات ذات الحظوة في المجتمع.

١٩٦٧ **هوبسون ضد هانسن:** تميز اختبارات القابلية الجماعية ضد أطفال الأقليات وعليه لا يمكن استخدام هذه الاختبارات لتوزيع الطلاب على "مسارات" مختلفة للقدرة.

١٩٧٠ **ديانا ضد مجلس التعليم العام:** لا يمكن استخدام إجراءات الاختبار التقليدية لتصنيف الأطفال من أصل مكسيكي على أنهم "معاقون يمكن تعليمهم" بل يجب استخدام إجراءات خاصة (كأساليب التقييم المعتمدة على لغتين) لاختبار أطفال الأقليات.

١٩٧١ **غريغس وغيره ضد شركة كهرباء ديوك:** يمنع استخدام إجراءات التوظيف التي تهدف لاستبعاد الزنوج والتي لا يمكن أن تبرر علاقتها بالأداء الوظيفي.

١٩٧٢ **غوادالوبه ضد مقاطعة تيمب الإبتدائية:** يجب أن يتم اختبار الأطفال بلغتهم الأولى ويجب، بناءً على ذلك، حذف عدد من الأسئلة غير العادلة من اختبارات الذكاء التي تقرر انضمام الأطفال إلى صفوف "المعاقين الممكن تعليمهم".

١٩٧٣ **الولايات المتحدة ضد شركة كهرباء جورجيا:** تم تعزيز سلطة توصيات الـ EEOCC في تحديد صدق اختبارات التوظيف. ويجب على الشركات المختلفة أن تثبت أن إجراءاتها لاختيار الموظفين تقيس المهارات اللازمة لتأدية العمل.

١٩٧٥ **شركة الورق ألبرمال ضد مودي:** لا يمكن الاعتماد على تصنيفات المشرفين في اثبات صدق اختبارات التوظيف وحتى في حال كون الاختبار صادق ولكنه يؤثر بشكل سلبي على توظيف من ينتمون إلى جماعات بشرية معينة يتعين على المؤسسة بذل قصارى جهدها لإيجاد وسيلة اختيار أقل تحيزاً.

١٩٧٦ **واشنطن ضد دايفيس:** تم توسيع قاعدة المعايير التي يصادق من خلالها على اختبارات اختيار الموظفين لتضم برامج تطوير الأداء التدريبية المقدمة أثناء الخدمة.

١٩٨٠ **لاري ف. ضد رايلز:** إن اختبارات الذكاء الفردية متحيزة ضد السود وعلى إدارة ولاية كاليفورنيا التعليمية أن تتوقف عن استخدام هذه الاختبارات في عمليات التشخيص والتصنيف التعليمية.

١٩٨٠ جمعية الآباء الناشطين في التعليم الخاص (PASE) ضد هانون: إن استخدام اختبارات WISCو -R WISC واختبار ستانفورد – بينيه مع مراعاة المعايير المنصوص عليها قانونيا لتحديد البرنامج التعليمي المناسب لطفل ما لا يميز ضد الأطفال السود.

١٩٨١ ديبرا ب. ضد تيرلينغتون: وقد أكد هذا الحكم حكماً سابقاً صدر عن محكمة فيدرالية يفيد بأن برنامج اختبارات ولاية فلوريدا لاختبار الحد الأدنى من الكفاءة هو برنامج غير دستوري لأنه يخالف بند الحماية المتكافئة التي ينص عليها الدستور.

١٩٨٥ NAACP في ولاية جورجيا ضد ولاية جورجيا: لا تميز اختبارات الذكاء ضد الأطفال السود ولا يشكل وجود أعداد كبيرة من الأطفال السود ضمن صفوف "المعاقين الممكن تعليمهم" دليلاً على التمييز العنصري في هذه الاختبارات.

١٩٨٨ واتسون ضد بنك وتروست فورت وورث: يمكن المصادقة على إجراءات التوظيف الشخصية كالمقابلات ولكن يمكن للموظفين الإدعاء على أرباب العمل على أساس "التأثير السلبي" الناتج عن الترقيات المبنية على المقابلات.

١٩٨٩ شركة واردس كوف للتغليف ضد أنتونيو: عكست فكرة أساسية في قضية غريغس ضد شركة كهرباء ديوك إذ حملت الموظفين أنفسهم مسؤولية إثبات عدم صدق أو ثبات الاختبارات النفسية المستخدمة في قرارات الترقية.

ولقد كان التحيز في الاختبارات النفسية كأدوات للاختيار والتنسيب في المؤسسات السبب الذي استندت إليه العديد من القضايا القانونية في العقود الأربعة الأخيرة (انظر الجدول ١–٣). ومن أول القضايا هي قضية مايرت ضد شركة موتورولا والتي استخدمت في اختيارها للموظفين اختباراً تم قياسه على عينة اختبار مؤلفة من البيض فقط وبناءً عليه حكمت المحكمة بأن الاختبار منحاز ضد الفئات المحرومة حضارياً. ونتيجةً لذلك أصدر الكونغرس الأمريكي قانون الحقوق المدنية لعام ١٩٦٤ والذي يمنع البند الثامن منه التمييز على أساس العرق واللون والأصل والجنس والدين. وقد وضع هذا القانون تحت

الاختبار في قضية غريغس ضد شركة كهرباء ديوك التي رفعت للمحكمة العليا إذ شككت هذه القضية بإجراءات شركة كهرباء ديوك في اختيار الموظفين والتي تشترط شهادة ثانوية بالإضافة إلى حد أدنى من الدرجات في اختبارووندرليك للموظفين واختبار الإسيعاب الميكانيكي. وقد أتى حكم المحكمة ليمنع استخدام أي إجراء توظيفي ليس له علاقة بالأداء الوظيفي يهدف إلى استبعاد الأفارقة الأمريكيين. ورغم اعتراف المحكمة العليا بفائدة الاختبارات النفسية في عملية اختيار الموظفين إلا أنها أكدت على أن هذه الاختبارات لا يمكن أن تعطى سلطة الإدارة إلا في حالات ارتباطها المباشر بقياس الأداء الوظيفي.

وقد هدفت قرارات المحكمة العليا في قضية موتورولا وقضية الولايات المتحدة ضد شركة كهرباء جورجيا وقضية شركة ألبرمار للورق ضد مودي إلى مطالبة أرباب العمل توضيح وتوثيق العلاقة بين المهارات التي تقيسها الاختبارات المختارة وبين الوظائف الشاغرة. ولكن، وكما يوضح قرار المحكمة في قضية واشنطون ضد دايفيس، لا يشترط في الاختبارات أن يصادق عليها بمقارنتها مع الأداء الوظيفي الفعلي فالأداء ضمن برنامج تدريبي مناسب هو معيار مناسب للمصادقة على اختبار مصمم لاختيار الموظفين.

وقد أدى الفشل في تطبيق قانون الحقوق المدنية لعام ١٩٦٤ في مواقف متعلقة بالتوظيف إلى صدور قانون تكافؤ الفرص الوظيفية لعام ١٩٧٢ وإلى تشكيل مجلس تنسيق تكافؤ الفرص الوظيفية (EEOCC). وقد ساهمت وثيقة التوصيات الموحدة لإجراء ات اختيار الموظفين والتي وضعها مجلس EEOCC في التعريف بالإجراءات التي يتوجب على أرباب العمل والمنظمات العمالية ووكالات التوظيف استخدامها لتوضيح صدق إجراءات اختيارهم للموظفين والتي يمكن أن تكون متحيزة ضد أعضاء أي جماعة من عرق أو لون أو جنس ما. إضافة إلى ذلك، وكما جاء في الحكم في قضية شركة ورق البمارل ضد مودي، فحتى عندما

يكون إجراء الاختبار تم إثبات صدقه ولكن له تأثير سلبي على توظيف جماعات بشرية ما يتعين على المؤسسة أن تحاول إيجاد إجراء اختيار آخر أقل تحيزاً. ويعرّف مبدأ "التأثير المضاد (السلبي)" قانوناً وفق قاعدة الأربعة أخماس على النحو التالي: تعد إنخفاض نسبة إختيار مجموعة بشرية ما عن أربعة أخماس من نسبة المجموعة ذات نسبة الإختبار الأعلى دليلاً على وجود حالة من التأثير المضاد.

وقد أدت إلتباسات متعلقة بمعنى المعايير التي لها علاقة بالعمل وما يعتبر دليلاً مقبولاً على صدق اختبار ما إلى تعديلات في توصيات الـ EEOCC لاختيار الموظفين عام ١٩٧٨، وطالبت التوصيات المعدلة كما كان حالها قبل التعديل أرباب العمل بتبرير استخدام اختبارات وإجراءات إختيار تستبعد أعداداً نسبها غير مقبولة من المنتمين إلى الأقليات أو من النساء. ووضحت التعديلات إلى حدٍ ما أساليب الادّعاء التي يمكن أن يستخدمها أرباب العمل لإثبات صدق إجراءات الاختيار التي يستخدمونها. واستناداً إلى الحكم في هذه القضية وغيرها من القضايا (مثل قضية شركة وواردس كوفر للتغليف ضد أنتونيو) تم تشريع قانون الحقوق المدنية لعام ١٩٩١ والذي (١) وضح شرط كون عبء إثبات صدق الاختبار يقع على عاتق رب العمل. (٢) ومنع بالقانون استخدام الدرجات الفاصلة التفاضلية التي تم تحديدها على أساس العرق أو الجنس أو الخلفية العرقية. وعليه فقد تم منع أي شكل من أشكال "تعديل الدرجات" المبني على أساس العرق أو اللون أو الدين أو الجنس أو القومية بموجب الفقرات ١٠٢-١٦٦ من الفقرة ١٠٦ من القانون العام. كما منع القانون ما يسمى بالمعايرة العرقية وهي إجراء تحدد فيه معايير الاختبارات لكل عرق بمعزل عن الأعراق الأخرى وتقارن درجات المتقدمين مع أولئك من الذين ينتمون إلى نفس المجموعة البشرية وذلك قبل إتخاذ قرارات توظيف بشأنهم.

القضايا القانونية في مجال الاختبارات التعليمية

تعرضت الاختبارات في مجال التعليم إلى انتقادٍ حاد، شأنها كشأن اختبارات التوظيف، بسبب تحيزها ضد الأقليات والنساء على وجه الخصوص. ولعل أهم القضايا المتعلقة بالتحيز ضد الإناث هي تلك المتعلقة باختبار القابلية المدرسية (التقدير في الوقت الحاضر) SAT Scholastic uptitude (assessment now) الذي يتفوق فيه عادةً الذكور. ونظراً لأن المنح الجامعية تعتمد بشكل أساسي على نتائج اختبار SAT، فإن أعداداً أكبر من الذكور يحصلون على هذه المنح (شيا، ١٩٩٤). وتبقى العلاقة بين تفاوت معدلات درجات الذكور والإناث في اختبار SAT وغيره من اختبارات القبول الجامعي (وبخاصة في الرياضيات) وطبيعة الاختبارات بحد ذاتها علاقة غير واضحة ومحفوفة بالجدل، تلعب المدارس والعوامل البيولوجية وغيرها من عوامل البيئة المحيطة بالطلاب دوراً فيها. فقد أثبتت الدراسات أنه لا يوجد أي مجال أكاديمي يتفوق فيه الذكور بشكل ملحوظ على الإناث (وو، ١٩٩٧) إلا أن نتائج اختبارات SAT تشير إلى عكس ذلك.

وبالإضافة إلى الاختلافات بين الذكور والإناث في نتائج اختبارات SAT، تتباين نتائج الأفراد من الفئات العرقية المختلفة إذ يحرز الأمريكيون من أصل أفريقي نتائج أدنى في قسمي الاختبار المتعلقين باللغة (SAT-V) والرياضيات (SAT-M) بشكل ملحوظ من أقرانهم البيض. إلا أن درجة كون هذه الاختلافات نتيجة لتحيز في الاختبار ذاته هي قضية لا تزال موضع للنقاش وبدون حل واضح. ومن القضايا الأخرى المتعلقة باختبار SAT مطالبة الطلاب الرياضيين بمعدل ٧٠٠ درجة في اختبار SAT كحد أدنى أو ١٧ درجة في اختبار الجامعات الأمريكي (ACT) ومعدل نهائي لا يقل عن ٢,٥. ومما أثار الكثير من الجدل وأدى أحياناً إلى دعاوى قضائية استخدام اختبارات الذكاء في تحديد مَن مِن طلاب المدارس يتوجب إلحاقه بفصول التعليم الخاص.

لطالما قامت السلطات التعليمية في الولايات المتحدة بالاعتماد على نتائج طلاب المدارس في اختبارات الذكاء المختلفة (الجماعية والفردية) في توزيع هؤلاء الطلاب على فصول التربية الخاصة. وفي واقع الأمر، يمكن القول أن هذا التوزيع هو أحد الأهداف الرئيسة لاستخدام اختبارات الذكاء في القرن العشرين. وتكمن المشكلة، التي لفتت النظر إليها عملية الدمج بين الأعراق والأجناس في المدارس، في وجود أعداد كبيرة (وغير متناسبة مع نسب المجموعات البشرية الأخرى) من الأفارقة الأمريكيين في عداد هذه الفصول الخاصة بسبب تدني معدلات درجاتهم في اختبارات الذكاء. وقد جاء قرار المحكمة في واحدة من أول القضايا التي أثارت هذا الموضوع (قضية هوبسون ضد هانسون) ليمنع استخدام اختبارات الذكاء الجماعية المتحيزة ضد الأفراد المنتمين إلى الأقليات في توزيع طلاب المدارس على مسارات مختلفة للقدرة. وفي قضيةديانا ضد مجلس التعليم العام صدر حكم بعدم جواز استخدام إجراءات الاختبار التقليدية لتصنيف الأطفال من أصل مكسيكي في ولاية كاليفورنيا على أنهم "معاقون يمكن تعليمهم" بل يجب اختبار أطفال الأقليات بلغتهم الأم. وقد أكدت قضية غوادالويه ضد مقاطعة تيمب الإبتدائية على أنه يجب أن يتم اختبار الأطفال بلغتهم الأولى كما طالبت المحكمة بحذف عدد من الأسئلة غير العادلة من اختبارات الذكاء وأن يكون الحد الأدنى من الدرجات الذي يقرر إلتحاق الطلاب بفصول التعليم الخاص أدنى بدرجتين من المعدل الوسطي للاختبار وأن تستخدم مقاييس للسلوك التكيفي في تقرير ما إذا كان الطفل متخلفاً عقلياً أم لا.

وقد صدر حظر في قضية لاري ب. ضد رايلز على استخدام اختبارات الذكاء في إلحاق الأطفال بفصول التعليم الخاص في مدارس ولاية كاليفورنيا الحكومية، إلا أن حكماً مغايراً صدر في قضية جمعية الآباء الناشطين في

التعليم الخاص (PASE) ضد هانون حيث قررت المحكمة أن "استخدام اختبارات WISCو WISC-R واختبار ستانفورد – بينيه مع مراعاة المعايير المنصوص عليها قانونيا لتحديد البرنامج التعليمي المناسب لطفل ما... لا يميز ضد الأطفال السود" (PASE ضد هانون، ١٩٨٠، ص. ٨٨٣). وصدر حكم مشابه – أن اختبارات الذكاء لا تميز ضد الأطفال السود – في قضية NAACP في ولاية جورجيا ضد ولاية جورجيا. وبناءً عليه تم رفع الحظر على استخدام اختبارات الذكاء في ولاية كاليفورنيا بحجة أن هذا غير عادل في حق الآباء من الأفارقة الأمريكيين الراغبين في إلحاق أطفالهم المعاقين عقلياً بفصول التعليم الخاص (بريدامايير، ١٩٩١).

هذا وتستمر الأسئلة والتساؤلات في المستقبل حول الطبيعة العادلة أو التحيزية لاختبارات اختيار الموظفين وقرارت اختيار وتوزيع الطلاب على فصول القدرات المختلفة في مجال التعليم. ومن الجدير بالذكر أن هنالك وجوه اقتصادية وتعليمية وسياسية وعلمية واجتماعية متعلقة بهذه القضايا وغالباً ما تفشل قرارات المحكمة في هذه القضايا في حل المشكلة. ومن المؤكد أن استخدام الاختبارات لن يتوقف إلا أنه، ولربما، ستصبح الاختبارات أكثر حساسية تجاه الإختلافات بين المجموعات البشرية المختلفة أو ستتخذ أشكالاً جديدة. ويجب على مستخدمي الاختبارات أن يكونوا على دراية بأن أهداف ومخططات الأفراد تختلف بإختلافهم. وإنه غالباً ما تكون الحلول الوسط هي خير الحلول لمثل هذا النوع من النزاعات الاجتماعية، وعلى الرغم من أن الإتفاق بين المتنازعين هو بلاشك إتفاق مؤقت إلا أنه ربما يبقى أفضل من السعي لحل على أساس كل شيء أو لاشيء إذ يؤدي هذا إلى تطويل وتعقيد عملية الصراع.

الخلاصة

يركز هذا الكتاب على تطوير وتطبيق اختبارات القدرات الإدراكية والمهارية. ومن أهم الاختبارات التي يهتم بها الكتاب الاختبارات التحصيلية واختبارات الذكاء واختبارات الاستعدادات الخاصة. كما يهتم الكتاب بالمشاريع والمنتجات وغيرها من أشكال الأداء التي لا تختبر بالشكل التقليدي إلا أنها تعمل على تقييم المهارات في مجالات إدراكية ونفسية – حركية.

ويعود استخدام الاختبارات والامتحانات بغرض اختيار موظفي المؤسسات الحكومية إلى أكثر من ٤٠٠٠ عام، إلا أن الاختبارات الشفهية والكتابية الرسمية لم تستخدم في الجامعات إلا منذ عدة مئات من الأعوام. ولم يتم الاتجاه إلى التعامل مع الاختبارات النفسية بطريقة علمية مستندة إلى الأبحاث إلا في النصف الثاني من القرن التاسع عشر. وكان فرانسيس غالون وألفرد بينيه وغيرهما من علماء النفس من رواد البحث في مجال قياس الذكاء. وعلى حين أثبتت اختبارات غالتون السايكومترية أنها غير عملية، استخدمت مقاييس ستانفورد – بينيه وما أتى بعدها بشكل واسع في تشخيص حالات التخلف العقلي وقياس القدرة العقلية بمختلف درجاتها. كما تم استخدام اختبارات الذكاء الجماعية مثل اختباري الجيش ألفا وبيتا بشكل أوسع من استخدام مقاييس الذكاء مثل ستانفورد – بينيه وشبيهاتها.

كما شهدت السنوات الأولى من القرن العشرين بالإضافة إلى اختبارات الذكاء الفردية والجماعية المقننة ظهور اختبارات مقننة للتحصيل والاستعدادات الخاصة. وقد استخدمت اختبارات التحصيل في المدارس والجامعات كمقاييس لما قد أحرزه الطالب في مادة أكاديمية أو في مقرر تعليمي ما، كما استخدمت

اختبارات الاستعدادات الخاصة مثل الأعمال المكتبية والمهارات النفسية - الحركية والميكانيكية في مجال اختيار الموظفين بشكل خاص.

وقد أدى النمو الذي شهدته "صناعة" الاختبارات النفسية والتعليمية منذ العشرينيات إلى تأسيس العديد من دور النشر المسوقة لمواد الاختبار مثل خدمة الاختبارات التعليمية (وهي أكبر هذه الدور) والدار الوطنية لأنظمة الحاسب ومكتب اختبارات كاليفورنيا وشركة علم النفس المتحدة والبرنامج الأمريكي للاختبارات الجامعية ودار نشر ريفرسايد (راجع الملحق ب والملحق ج). وتسوق دور النشر هذه وغيرها من المؤسسات العديد من أدوات المقاييس السايكومترية في مجال كل من الإدراك والشعور والمجال الحسي - الحركي عند الأفراد. وقد استخدمت هذه الاختبارات بشكل واسع وغالباً ما كان لها تأثيرات مثيرة للجدل على البرامج التعليمية والتوظيفية والإرشادية والعلاجية. كما ترعى وتعد بعض المنظمات الحكومية والخاصة كاتحاد الجامعات الطبية الأمريكية، والمجلس الوطني لاختبارات المحامين، والمفوضية الأمريكية للوظائف الحكومية، اختباراتٍ غير تجارية. ومن ضمن الأعداد الكبيرة من المصادر المتوفرة عن الاختبارات المنشورة وغير المنشورة المصادر التالية: كتاب الاختبارات المنشورة، وكتاب الاختبارات، والكتب السنوية للمقاييس العقلية، وكتاب مقالات نقدية عن الاختبارات بالإضافة إلى عدد من قواعد البيانات الالكترونية والمقالات في المجلات التخصصيةوالاختبارات بحد ذاتها ودلائل المستخدم التي ترفق بالاختبارات. وعلى الرغم من أن معظم الاختبارات المذكورة في المصادر أعلاه هي اختبارات تجارية منشورة، هنالك نسبة كبيرة من الاختبارات غير المنشورة والتي تطورها الهيئات الحكومية والشركات الخاصة بهدف استخدامها داخلياً.

كما يلتزم مستخدمو الاختبارات المحترفين ودور النشر ذات السمعة الطيبة بمقاييس أخلاقية حددتها منظمات رسمية كالاتحاد الأمريكي لعلم النفس والاتحاد الأمريكي للأبحاث التعليمية والاتحاد الأمريكي للتوظيف والإرشاد. ولعل من أهم المصادر الجديرة بالذكر هنا المصادر التالية: مقاييس الاختبارات النفسية والتعليمية، وتوصيات للاختبارات المحوسبة وتفسير نتائجها، و"دستور الممارسات الاختبارية العادلة في مجال التعليم والمبادئ الأخلاقية لعلماء النفس ودستور سلوكهم". وقد أكدت جميع هذه المصادر على ضرورة تسويق واستخدام الاختبارات ضمن درجة من الحساسية للخصائص السايكومترية لهذه الاختبارات (من حيث الصدق والثبات وإجراءات القياس... وغيرها)، ولكفاءات القائمين عليها، ولمصلحة الممتحنين والمصلحة العامة. كما يجب أن لا يتم اختبار أي فرد دون موافقته أو موافقة وكيله القانوني، ويجب التعامل مع النتائج التي يتوصل إليها الاختبار ضمن مبادئ السرية والخصوصية المنصوص عليها، وأخيراً يتم التعبير عن الحالات التي تشخصها الاختبارات بطريقة تخلو من التشهير والتجريح.

وقد ركزت الدعاوى القضائية في مجال استخدامات الاختبارات النفسية والتعليمية في سياقي التعليم والتوظيف على نزاعات متعلقة باختيار وتوزيع الموظفين ضمن المؤسسات المختلفة. وقد تناول هذا الفصل من الكتاب بعض الدعاوى القضائية الهامة والتي كانت أسبابها في مجال التوظيف تعود إلى إدعاءات بفقدان العدل في سياسات اختيار الموظفين المنتمين إلى الأقليات، فيما كانت أسباب القضايا في مجال التعليم مبنية على إدعاءات بالتوزيع غير العادل لأطفال الأقليات على صفوف التعليم الخاص. وغالباً ما كانت اختبارات الذكاء هي السبب الرئيسي للجدل والنزاع.

مصطلحات للمراجعة

راجع معاني المصطلحات المدرجة أدناه والتي تم استخدامها في هذا الفصل. الرجاء مراجعة فهرس التعريفات أو المعجم في حال وجود أي التباس حول معاني المصطلحات.

القدرة	اختبار مبني على مرجعية المعيار
الاختبار التحصيلي	العدالة
التأثير السلبي (المضاد)	السؤال محدد البدائل
المقاييس العاطفية	التقييم التكويني
المعيار الموافق للعمر	قاعدة الأربعة أخماس
اختبار الاستعداد	الاختبار الجماعي
العمر القاعدي	الاختبار الفردي
العمر السقفي	الموافقة المطلعة
القدرات الإدراكية	الذكاء
تفسير (تحليل) نتائج الاختبارات	تشخيص الحالة دون تجريح
عن طريق الحاسب (CBTI)	اختبار الاتقان
العمر العقلي	المهارات الحركية – النفسية
اختبار مبني على مرجعية المحك	المعايرة المبنية على العرق
المعايير	معدل الذكاء النسبي IQ
الاختبار الموضوعي	اختبار السرعة
السؤال المفتوح	شكل الدليل اللولبي
اختبارات الأداء	الاختبار القياسي
اختبار القوة	التقييم التراكمي
المعلومات المحصنة من الإفشاء	القدرات اللغوية

المراجع المقترحة:

Bowman, M. L. (1989). Testing individual differences in Ancient China. *American Psychologist, 44,* 576–578.

Brown, D. C. (1994). *Subgroup norming: Legitimate testing practice or reverse discrimination?* Washington, DC: American Psychological Association.

Doyle, K. O., Jr. (1974). Theory and practice of ability testing in Ancient Greece. *Journal of the History of the Behavioral Sciences, 10,* 202–212.

Forrest, D. W. (1974). *Francis Galton: The life and work of a Victorian genius.* London: Elek.

Goldstein, G., & Hersen, M. (1990). Historical perspectives. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (2nd ed., pp. 3–17). New York: Pergamon Press.

Robertson, G. J. (1992). Psychological tests: Development, publication, and distribution. In M. Zeidner & R. Most (Eds.), *Psychological testing: An inside view* (pp. 159–214). Palo Alto, CA: Consulting Psychologists Press.

Rosenbach, J. H., & Rusch, R. R. (1991). IQ and achievement: 1930s to 1980s. *Psychology in the Schools, 28* (4), 304–309.

Sandoval, J., & Irvin, M. G. (1990). Legal and ethical issues in the assessment of children. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence and achievement* (pp. 86–104). New York: Guilford Press.

Sokal, M. M. (1987). *Psychological testing and American society 1890–1930.* New Brunswick, NJ: Rutgers University Press.

Thorndike, R. M., & Lohman, D. F. (1990). *A century of ability testing.* Chicago: Riverside.

Wyche, L. G., & Novick, M. R. (1992). Standards for educational and psychological testing: The issue of testing bias from the perspective of school psychology and psychometrics. In A. K. H. Burlew, W. C. Banks, H. P. McAdoo, & D. A. ya Azibo (Eds.), *African American psychology: Theory, research, and practice* (pp.229-235). Newbury Park, CA: Sage.



المبادئ والطرق السايكومترية

لعل أحد أهم خصائص الإدراك البشري القدرة على تشريح وتحليل التجارب وعلى تفصيل أوجه وعناصر الإدراكات الحسية والمفاهيم بغرض التعامل معها بشكل فعال. إلا أن المشكلة في تحليل الأشياء إلى عناصرها أو مكوناتها المختلفة تكمن في اختلاف طريقة التحليل ونتائجه من شخصٍ لآخر. ويتطلب التواصل الفعال والعمل التعاوني بين الأفراد نوعاً من الإجماع على الإجراءات والمتغيرات التي نحلل بواسطتها تجاربنا الحياتية. وحالما نتمكن من وضع نظامٍ موحد للقواعد التي تنظم عمليات تحليل وتصنيف الظواهر التي نراقبها ونفكر ونتحكم بها، سيصبح استخدام عملياتنا العقلية على أكمل وجه ممكن. وسيمكننا هذا من أن نتحدث عن تجاربنا مع الآخرين وأن نستخدم هذه التجارب بطرق قديمة ومعروفة أوحديثة ومبتكرة. كما سيمكننا من أن نبيع هذه التجارب أو نشتري خلاصة تجارب الآخرين أو أن نجمع تجاربنا مع أشياء أخرى أو أن نتنبأ بالتجارب المقبلة. وإن إيجاد نظام للقياس يسهل كافة العمليات الآنفة الذكر.

القياس والإحصاء

القياس

يُعرّف القياس على أنه ربط الأشياء والأحداث بقيم عددية وهو قديمٌ قدم الإنسانية إن لم يكن أقدم[1]. ولقد كانت أولى أشكال القياس معنية بالخصائص الفيزيائية للأشياء مثل الطول والعرض والإتساع والزمن التي استخدمت في مجالات عملية وتطبيقية مثل الانتاج والبناء والتجارة وتوزيع الموارد. فعلى سبيل المثال، ساهمت الحسابات الفلكية التي قام بها البابليون وغيرهم من الشعوب القديمة في التخطيط لزراعة المحاصيل في الأوقات الأكثر خصوبة. كما طور الإغريق نظاماً رياضياً تجاوز حدود الاستخدامات العملية للقياس وشكل أساساً للعلوم النظرية والتطبيقية. ومن أحد العوامل الأخلاقية واللاأخلاقية التي ساهمت في التطورات التي حققها العالم القديم هو نظام العبودية الذي حرر بعض اليونانيين الأذكياء من تعب العمل لكسب المعيشة وبالتالي منحهم الوقت للتفكير في أمور أبعد مدىً من اعتبارات الحياة العملية اليومية لديهم.

يمكن للعلم، إلى حدٍ ما، أن يستمر دون تحديد المتغيرات بكميات (أو أعداد) معينة، إذ اعتمد شكل التفكير العلمي الذي أستخدمه أرسطو كلياً على الوصف والتصنيف اللاعدديين، كما استمرت العلوم وبالذات علم الأحياء لقرون بعد أرسطو في الوصف والتصنيف ووصلت هذه الطريقة إلى قمة تطورها في نظام لينييوس للتصنيف العلمي. يعتمد هذا النظام على وضع الأشياء ذات الخصائص المتشابهة ضمن الفئة ذاتها، وتتصف هذه الطريقة بكونها طبقية كما يوضح المقياس التاريخي العرقي الذي يبدأ بالمملكة ثم الشعبة أو الفصيلة ثم الصف ثم الطبقة ثم العائلة ثم النوع ثم الصنف ثم العرق. وتتشابه الأحياء المنتمية إلى

---

(١) تمتلك بعض الحيوانات والطيور قدرات حسابية بسيطة تمكنها من معرفة ما إذا نقص عدد بيوضها أم لا وتسمى هذه القدرة "subitism".

الصف ذاته أكثر من تشابه الأحياء المنتمية إلى نفس الشعبة وتتشابه تلك التي تنتمي إلى الصنف ذاته أكثر من تلك المنتمية إلى نفس النوع. وبعد أن تصنف النباتات أو الحيوانات وفق المقياس التاريخي العرقي وتعطى اسماً مناسباً (محدداً من حيث الصنف)،يتم التنبؤ بردود أفعالها وسلوكها بمقارنتها مع غيرها من الكائنات المنتمية إلى نفس الفئة. ويبقى الاتجاه الأرسطي (نسبةً إلى أرسطو) اتجاهاً محدوداً بالوصف لا يوفر أي تفسيرات.

وعلى شكل مغاير للأبحاث الأرسطي الوصفي/التصنيفي للعلوم، يتعامل الاتجاه الغاليلي (نسبة إلى غاليليو) الوظائفي مع المتغيرات والمبادئ التي تنطبق على الأشياء والأحداث ومع صياغة هذه المبادئ بأسلوب قانوني ورياضي. ومن الأمثلة على هذه المعادلات والصيغ هي القانون $s = {}^{1}/_{2}\, gt^{2}$ أو المسافة التي يقطعها جسم يسقط سقوطاً حراً في الفراغ يساوي نصف ثابت تسارع الجاذبية (g) مضروباً بمربع الوقت محسوباً بالثانية (t).

الرياضيات وعلم النفس

على عكس العلوم الفيزيائية، تميزت الجهود العلمية في مجال علم النفس وغيرها من العلوم الاجتماعية بكونها وصفية لم تتم صياغة مبادئها وقوانينها ضمن معادلات رياضية. ويستثنى من هذا التعميم القانون النفسي الفيزيائي الذي اقترحه غوستاف فنتشرمنذ أكثر من مئة عام وهو: $S = k\log R + c$ أو مقدارالإحساس(S) هو تابع لوغاريتمي لشدة عامل الإثارة (R).

إن مصطلح القياس النفسي، من وجهة نظر الفيلسوف إيمانويل كانت، هو تناقضٌ في المصطلحات بحد ذاته، إذ يرى كانت أن القياس الموضوعي هو من خصائص العلوم وأن علم النفس هو علم ذاتي غير موضوعي يقتصر عمله على الظواهر الذاتية التي لا يمكن قياسها. إلا أن الدراسات التي قام بها

فينتشروبينيه وغيرهما في نهاية القرن التاسع عشر وبداية القرن العشرين أثبتت أنه يمكن التعبير عددياً وكمياً عن التجارب الواعية والسلوك البشري الملاحظ. وعلى الرغم من أن هذا القياس العددي ليس على نفس الدرجة من الدقة الموجودة في الأحداث الفيزيائية، إلا أنه يبقى نوعاً من القياس.

يُعرف القياس النفسي كما يستخدم في مجالات الاختبارات النفسية والتعليمية بأنه "عملية إعطاء قيم عددية (مثل درجات الاختبارات) للأشخاص بطريقة تسمح لخصائص هذه الأعداد أن تعكس وبشكل دقيق صفات هؤلاء الأشخاص" (مورفي وديفيدشوفر، ١٩٩٤، ص. ٣٤). ويجب الملاحظة هنا أن هذا التعريف لا يدّعي بأن علم النفس قادر على قياس كافة أفكار الفرد ومشاعره وأنماط سلوكه إذ أن أهداف القياس النفسي لا تطمح إلى كل هذا بل تكتفي برسم لوحاتٍ تجريدية تعبّر عن الاختلافات في ردور أفعال الأفراد الداخلية والخارجية.

وكما ذكرنا في الفصل الأول، بدأ التركيز العلمي على قياس الاختلافات بين الأفراد في أعمال فرانسيس غالتون ولا زال هذا التركيز هدفاً أساساً للقائمين على عمليات القياس النفسي. ويجب على المقاييس النفسية لكي تحقق أهداف التنبؤ بطبيعة الأشخاص والسيطرة على بعض طرق سلوكهم أن تستخدم مقاييس ثابتة، ولتحقيق هذا الثبات يجب أن تكون أدوات القياس مستقرة تحت الظروف ذاتها (هذا في حال أن بيئة الفرد لم تتغير بشكل ملحوظ). وعليه فإن المقاييس الثابتة هي تلك التي تتطلب تطبيق إجراءات القياس ذاتها، أي تتطلب قياس الاختبارات. وإضافة إلى ذلك يشترط في المقاييس أن تكون صادقة: أي يجب أن توفر قيم عددية صحيحة تقيس الخاصية أو المتغير المراد قياسه.

"العملياتية" Operationalism

تتصف جميع أنظمة القياس بكونها افتراضية بعض الشيء، إذ أن المؤشر الذي تقع عليه المقاييس لا وجود له كصفات ملموسة في الواقع، وإن الأرقام الموجودة على المؤشر هي فقط طريقة عملية وسهلة لتمثيل نوعية أو شدة الشيء أو الحدث موضع القياس. وإن المتغيرات التي يتم قياسها ثم تحديد قيم عددية موافقة لها ضمن ظروف معينة أو ضمن إطار زمني معين هي ليست إلا ابتداعات خيالية ذات فائدة عملية تساعدنا في تبسيط ملاحظاتنا وتعاملاتنا وفهمنا لأنفسنا وللعالم من حولنا. ويسمى الميل إلى التصرف على أساس أن هذه المفاهيم والمجردات هي أشياء ملموسة لها وجود في الواقع وتدعى Reification (أي تحويل المجردات إلى أشياء مادية ملموسة)، وقد كان هذا المحفِّز الأساسي لنظرية بريدجمان العملياتية.

تركز نظرية العملياتية على أن إلتباس معاني المتغيرات أو المفاهيم كمفهوم القوة مثلاً يمكن تفاديه عن طريق تعريف هذه المفاهيم من حيث العمليات الضرورية والتي يمكن ملاحظتها التي تؤدي إلى شعورنا أو مرورنا بتجربة تتجلى فيها هذه المفاهيم. وببساطة يعد التعريف العملياتي تعريفاً يدلنا على ما يتوجب علينا فعله أو على الإجراءات التي يجب أن نتبعها بغرض تحفيز الشعور المراد تعريفه في داخلنا. فعلى سبيل المثال، يتطلب التعريف العملياتي للطول مثلاً إعطاء شخصٍ ما تعليمات بكيفية قياس ما يسمى بالطول، ويتطلب التعريف العملياتي للقلق وصفاً دقيقاً لتصرفات الفرد القلق وللأحداث أو المحفزات التي أثارت الشعور بالقلق عند هذا الشخص.

وتساعد التعريفات العملياتية للمفاهيم في تسهيل عملية التواصل وتوحيد معاني المفاهيم. إلا أن المبالغة في استخدام هذه التعريفات قد تؤدي إلى عقم في عملية البحث العلمي، إذ أن المفاهيم العلمية لا يمكن تعريفها دائماً بطريقة

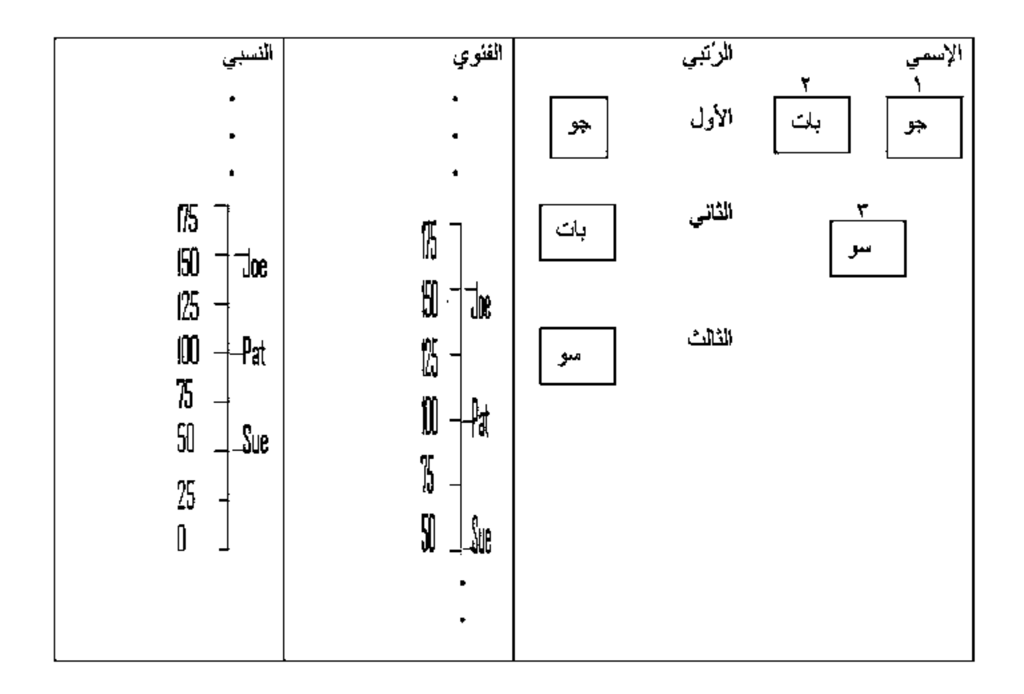
عملياتية، على الأقل ليس فوراً، وإن التأكيد الزائد على ربط المتغيرات بما يمكن ملاحظته ومراقبته وتعريف هذه المتغيرات على أساس الدلالة (أو الإشارة) يمكن أن يثبط التخمين المبدع والتفكير التجريدي. وقد أكدت الحركة العملياتية في علم النفس على أن مفاهيماً كالقلق والتعلم والإحباط لا تعني الشيء ذاته للأفراد المختلفين ولذلك فإن هذه المفاهيم لا تعبر عن التجارب ذاتها. وبما أن العلم الجيد يرتكز إلى توثيق الملاحظات والتنبؤات على أرض الواقع،يتوجب على العلماء أن يصفوا ما تمت ملاحظته أو ما قد حصل في موقف ما بأوضح شكل ممكن. ولعل أفضل طريقة لفعل هذا، وذلك عندما تكون الظاهرة موضع الدراسة يمكن تحديد مقدارها، هي بالإشارة بدقة متناهية إلى كيفية قياس المتغيرات المراد قياسها.

مؤشرات القياس Scales of Measurements

كما أشرنا سابقاً في هذا الفصل، إن أبسط تعريف للقياس هو الإشارة إلى الأحداث بواسطة الأرقام. واستناداً إلى هذا التعريف تعتبر الأرقام التي يحملها الرياضيون والتي تساعدنا على رصد ومتابعة أدائهم نوعاً من القياس. وبما أن جمع أو طرح أو تقسيم أو ضرب هذه الأرقام عملية لا تعني شيئاً، فإن هذا النوع من القياس يمثل أدنى درجات القياس أو ما يسمى بالمؤشر الإسمي. وبالنظر إلى الحقل الأول من الشكل ٢-١ يمكن القول أن جو وبات وسو ينتمون إلى فئات مختلفة من المتغير المحدد أعطيت الأرقام ١ و٢ و٣، وإن هذه الأرقام هي افتراضية تماماً ولا يمكن تحليلها بشكل ذو معنى إلا بإعتبارها أدوات مناسبة للتعرف على هؤلاء الأفراد.

الشكل ٢-١ أربع مؤشرات للقياس

مستويات القياس



يعتبر المقياس الرُتبي إلى حد ما مستوىً أعلى من مستوى القياس الإسمي لأنه يرتب ما يقاس من حيث الأول فالثاني فالثالث... إلخ. ومن الممكن أن تمثل هذه الأعداد الترتيبية تسلسل النجاح في مسابقة ما أو غيرها من المتغيرات التي ينال المشتركون فيها رتب مختلفة. ومن الملاحظ في الشكل ٢-١ أن جو أحتل المركز الأول على حين أحتل بات وسو المركزين الثاني والثالث على التوالي. وإن عملية الجمع بين هذه الأعداد بواسطة عمليات رياضية هي عملية لا معنى لها. وبما أن المسافة بين الأعداد غير معروفة فإنه لا يمكن القول بأن الفرق بين جو وبات لا يساوي الفرق بين بات وسو، ولذلك فإن كل ما يمكن الخلاص إليه في هذه الحالة هو أن جو أفضل من بات وبات أفضل من سو.

تصبح الأعداد أكثر دقة ووضوحاً عند إنتقالنا لمستوى القياس الفئوي، فكما نلاحظ في العمود الثالث من الشكل ٢-١، لا تتحدد مؤشرات القياس المقسمة بالتساوي بحدود عليا وحدود دنيا. وعلى عكس المقياس الرتبي، تتوافق الفوارق بين الأعداد الموجودة على المؤشر مع الفروقات المتساوية في أداء الممتحنين في المتغير موضع القياس. وبما أن معدل درجات جو هو ١٥٠ ومعدل درجات بات هو ١٠٠ ومعدل درجات سو هو ٥٠، فإن الفرق بين جو وبات يساوي الفرق بين بات و سو في الصفة التي تم قياسها. وتعد مقاييس الفهرنهايت والسيلسيوس لدرجات الحرارة أمثلة شائعة للقياس الفئوي. وفي مجال القياس النفسي يمكن إعتبار مؤشرات الدرجات المعيارية التي تعبر عن الأداء في الاختبارات الإدراكية مقاييساً فئوية.

وعلى حين يمكن طرح وجمع الأعداد الموجودة على المؤشر الفئوي، لا يسمح بإستخدام عمليتي القسمة والضرب في هذا المستوى من القياس إذ يتوجب على الأعداد أن توضع على مؤشر نسبي حتى يمكننا أن نقول أن معدل درجات بات (١٠٠) هو ضعف معدل درجات سو (٥٠) وأن معدل درجات جو (١٥٠) هو ثلاثة أضعاف معدل درجات سو.وكما يوضح العمود الرابع من الشكل ٢-١ فإن المؤشر النسبي له صفر مطلق أو أصل حقيقي تكون قيمة المتغير أو الصفة المقاسة فيه تساوي الصفر. وعليه، يمكن التعبير عن قيمتين عدديتين بعيدتين بمسافة محددة عن الصفر على مقياسي الطول (أمتار) والوزن (غرامات) كمعدل نسبة، كالقول مثلاً إن الدرجة ١٠٠ تمثل ضعف كمية الصفة التي تم قياسها بالدرجة ٥٠ وهي أيضاً تمثل ثلثي الدرجة ١٥٠. ومن الملاحظ أن السقف أو الحد الأعلى لمؤشر القياس النسبي هو لامحدود ولكن هذا لا يؤثر على خصائصه المتعلقة بالنسب. لاتُستخدم المؤشرات النسبية على نحو واسع في مجال القياس النفسي

باستثناء مقياس "سون" Sone لقياس علو الصوت ومقياس "مل" Mel لقياس طبقة الصوت.

الطرق الإحصائية

يقدم الجزءان الأول والثاني من هذه السلسة وبالتحديد الفصل الرابع من كتاب سلالم التقدير وقوائم الرصد (أيكين، ١٩٩٦) والفصلان ٤ و٥ من كتاب الاستبانات ومقاييس السمات (أيكين، ١٩٩٧) ملخصاً لابأس به عن طرق الإحصاء الأولية التي استخدمت لوصف عينات من المعلومات واستنتاج المؤشرات العامة لباقي السكان. وتعد معظم المفاهيم والإجراءات الإحصائية التي تناولتها الفصول المذكورة أعلاه مناسبة لتحليل البيانات الناتجة عن الاختبارات وغيرها من أدوات التقييم، كما توفر العديد من كتب الإحصاء الأولي (مثل: كايس، ١٩٩٦؛ باغانو، ١٩٩٤؛ شافيلسون، ١٩٩٦) ملخصات جيدة عن طرق الإحصاء البسيطة. وتتضمن أدلة المستخدم التي ترافق العديد من برامج الحاسب الإحصائية (مثل: نوروسيس، a١٩٩٢، b١٩٩٢، c١٩٩٢ على معلومات وصفية عن عمليات إحصائية متقدمة. ولن ندخل في تفاصيل الإحصاء هنا نظراً لوفرة المصادر حول هذا الموضوع، ولكن يبقى التعرف على إجراءات إحصائية مختارة لتحليل البيانات الناتجة عن الاختبارات النفسية وإعطاء لمحة عنها أمراً مفيداً وضرورياً.

إن أحد أول الخطوات المستخدمة في عملية تحليل بيانات الاختبارات النفسية هي التوزيع التكراري للدرجات التي حصل عليها الأفراد المختلفون، ويمكن تحقيق هذا من خلال العديد من برامج الحاسب كبرنامج A-I الموجود على القرص المرن المرفق بهذا الكتاب. ويعطينا هذا البرنامج تمثيلاً تصورياً

للتوزيع التكراري للنتائج عن طريق تحويل النتائج إلى مدرج تكراري أو مضلع للتوزيع التكراري. ويمكن فحص شكل هذه الرسوم البيانية بما في ذلك عدد الحد الأعلى للنقاط (المنوالات) ومدى تباعد أو اختلاف الدرجات وما إذا كان التوزيع متناظراً أم منحرفاً أو مسطحاً أم منحدراً. ويمكن لحساب الوسط الحسابي والوسيط والتباين والانحراف المعياري ومعاملات الالتواء والتفرطح عن طريق برنامج A-2 أن تكمّل المعلومات الناتجة عن طريق الفحص البصري للتوزيع التكراري.

ولعل معاملات الارتباط هي من أكثر الإحصائيات المتوفرة في الدليل المستخدم في الاختبارات النفسية بالإضافة إلى جداول الدرجات المعيارية والمئوية. وتوجد أنواع عديدة لمعاملات الإرتباط يمكن حسابها بواسطة برامج الفئة C الموجودة على القرص المرن المرفق بكتاب درجات التقدير وقوائم الرصد وبرامج الفئة E في كتاب الاستبانات وقوائم السمات. ويساعد برنامج A-3 الذي يحتويه القرص المرافق لهذا الكتاب في رسوم بيانية تبعثرية أو خطوط للارتداد الخطي لواحد من متغيرين بالمقارنة مع الآخر. كما يحسب برنامج A-4: (١) معامل الإرتباط المولد للعزوم بين متغيرين و(٢) الوسيط الحسابي والانحراف المعياري للمتغيرات و(٣) معادلة التراجع الخطي للتنبؤ بالمتغير التابع (Y) من المتغير المستقل (X) و(٤) عدد الأخطاء الإيجابية غير الحقيقية والأخطاء السلبية غير الحقيقية، التوقعات الصائبة وعدد حالات الرفض الصحيحة المرتبطة بحد للدرجات الفاصلة على المتغير (المعيار) التابع. ويمكن للبرنامج أن يساعد في دراسة صدق الاختبار في التنبؤ بمتغير معياري ما.

يمكن إجراء عمليات أخرى على مجموعة من معدلات الدرجات الناتجة عن عدد من الاختبارات مثل: تحليل الإنحدار المتعدد والتحليل العنقودي والتحليل

العاملي والتحليل التمييزي، ومن العمليات الأقل شيوعاً والتي قد تكون مفيدة عند وجود عدة متغيرات يمكن ذكر عملية تحليل الارتباط القانوني وتحليل المسار وتحليل البنية الكامنة. تقود النتيجة النهائية لتحليل الإنحدار المتعدد إلى معادلة إنحدارية يتم استخدامها في توقع درجات الأفراد في متغير تابع (معيار) بالاعتماد على درجاتهم في متغيرين مستقلين (أداوات التنبؤ) أو أكثر. وتتوفر عدة إجراءات لاختيار المتغيرات المستقلة التي سيتم استخدامها ضمن معادلة الإنحدار التعددي النهائية (كالاختيار الأمامي والحذف التراجعي والاختيارالمرحلي) (راجع نوروسيس، a١٩٩٢). ويتم إدخال جميع المتغيرات المستقلة (١-٣) بشكل متزامن في معادلة الإنحدار باستخدام برنامجA-5 الموجود على القرص المرافق لهذا الكتاب. كما يفيد برنامج A-5 في حساب التالي: قيم الانحدار القياسية وغير القياسية والأخطاء المعيارية لقيم الإنحدار والنتائج الإحصائية للاختبارات (t) لدلالات قيم الإنحدارومعامل الإرتباط المتعدد (R).

وعلى عكس تحليل الإنحدار المتعدد الذي يحتوي على متغير تابع واحد، يقوم تحليل الإرتباط القانوني على عدد من المتغيرات المستقلة والتابعة وذلك بهدف تحديد التركيب الخطي الأمثل لعدد من المتغيرات المستقلة الموزونة بغرض تحقيق الصحة الأمثل للتنبؤ بمجموعة من المتغيرات التابعة الموزونة عددياً. ومن الجدير بالذكر أن عملية تصنيف المتغيرات كمتغيرات مستقلة أو تابعةفي تحليل الإرتباط القانوني بين مجموعتين من المتغيرات عملية افتراضية تماماً.

تتطلب الإجراءات الإحصائية المتعددة المتغيرات كالتحليل العاملي وتحليل التمييزي جهداً حسابياً هائلاً، إلا أن مجموعة برامج الحاسب +SPSS/PC تنجز هذه العمليات بسهولة وسرعة (نوروسيس،b١٩٩٢)، وقد صممت إجراءات التحليل العاملي لتحديد العدد الأدنى من العوامل (الأبعاد) المطلوبة لتفسير

مجموعة معدلات درجات حصل عليها عدد nمن الأفراد في عدد من اختبارات m. وتعد الطرق التالية من أكثر طرق التحليل العاملي استخداماً: طرق المحور الرئيسي والحد الأدنى للمربعات غير الموزونة والاحتمالية القصوى وإجراءات ألفا وطرق التحليل العاملي للصور[٢]. وتنقسم عملية التحليل العاملي إلى مرحلتين، تتألف أولاهما من عملية استخلاص العوامل من مصفوفة بيانات الدرجات أومعاملات الإرتباط، بينما تنطوي المرحلة التالية على تدوير العوامل المستخلصة لتصبح بنى بسيطة يمكن تفسيرها بشكل ذو معنى. وتعد طرق فاريماكس وإيكواماكس وكواتريماكس أكثر طرق التدوير شيوعاً.

ويشتمل الجدول ٢-١ على قائمة لأحمال ثلاثة معاملات تم تدويرها استخلصت عن طريق تحليل معدلات درجات ٢٠٠٠ شخص على مقياس ويكسلر المعدل لذكاء البالغين (WAIS-R) وهذه الأحمال هي بالضرورة إرتباطات بين ١١ اختباراً فرعياً لمقياس WAIS-R مع معاملاتها، إذ ترتفع أهمية معامل ما في بنية الإختبار الفرعي بارتفاع أحماله. ويطلق على العامل رقم ١ اسم العامل اللغوي بسبب ارتفاع أحمال المعلومات والمفردات والإستيعاب والمتشابهات (وهي اختبارات فرعية لغوية) فيه. ونظراً لارتفاع أحمال إكمال الصور وترتيب الصور وتصميم المجسمات وتجميع الأجسام في العامل ٢، ونظراً لطبيعة هذه المهام يجب أن تطلق تسمية العامل الفراغي -الإدراكي - الحسي أو عامل الصور الفراغية على هذا العامل. وأخيراً، ترتفع أحمال المدى الرقمي والحساب والرموز العددية على العامل ٣، مما يعني أن اسم عامل المهارة العددية هو اسم مناسب لطبيعة هذا العامل.

---

(٢) وهي أحد طرق استخراج العوامل من المعلومات المعطاة. وتعتمد هذه الطريقة على استخراج العوامل من مصفوفة الإرتباط للمتغيرات المتنبأ بها وليس للمتغيرات الفعلية. (ملاحظة المترجم).

يهدف التحليل التمييزي لدرجات مجموعات من الأشخاص p في مهام معينة m إلى تحديد التركيبة الخطية للدرجات في الاختبارات m التي تساعدنا على التمييز بين المجموعات p. ويمكن مثلاً إجراء تحليل تمييزي لدرجات أفراد من أربع فئات عرقية (بيض، سود، آسيويون، إسبانيو الأصل) في بطارية اختبار للاستعدادات المتعددة وذلك بهدف تحديد التركيبة المرجحة الأمثل لمعدلات الدرجات القادرة على التنبؤ بأعضاء كل جماعة عرقية.

الجدول ٢-١

مصفوفة العامل المدور للاختبارات الفرعية لمقياس WAIS-R

الاختبار الفرعي	العامل ١	العامل ٢	العامل ٣
المعلومات	٠,٨١	٠,٢١	٠,٢٩
المدى الرقمي	٠,٢٧	٠,١١	٠,٨٣
المفردات	٠,٨٢	٠,٢٢	٠,٣٤
الحساب	٠,٤٩	٠,٢٨	٠,٥٩
الاستيعاب	٠,٧٩	٠,٢٣	٠,٢٦
المتشابهات	٠,٧٥	٠,٣٠	٠,٢٥
إكمال الصور	٠,٤٨	٠,٦٤	٠,١١
ترتيب الصور	٠,٥١	٠,٤٩	٠,١٣
تصميم المجسمات	٠,٢٦	٠,٧٢	٠,٣٨
تجميع الأجسام	٠,١٤	٠,٨٥	٠,١٩
الرمز العددي	٠,٢٣	٠,٣٧	٠,٦١

توفر عمليات التحليل التمييزي عدداً من المعلومات الإحصائية مثل معاملات الدالة التمييزية لعدة اختبارات وللدرجات التمييزية التي يتم حسابها من هذه المعاملات. ويساعد جدول تصنيفي ذو اتجاهين يبين عدد الأشخاص الذين تم تصنيفهم ضمن مجموعة معينة بشكل صحيح بناءً على درجاتهم الدالة التمييزية وعدد الأشخاص الذين تم تصنيفهم بشكلٍ خاطئ في إعطاء ملخص جيد عن فعالية درجات الاختبارات في التمييز بين المجموعات.

وكما هو الحال في تحليل العامل، يُعنى تحليل البنية الكامنة بتحديد درجة إمكانية تفسيرالعلاقات الملاحظة بين نتائج اختبارات p من خلال عدد من

المتغيرات الكامنة (أو غير الملاحظة). ويتعامل إجراء تحليل المسار الإحصائي مع العلاقات بين المتغيرات أيضاً، إلا أن التركيز في هذا الإجراء هو على تحليل الإرتباط بين مجموعة من المتغيرات بهدف تحديد أي من هذه المتغيرات تكون محتملة المساهمة في غيرها من المتغيرات أو مسببات لغيرها من المتغيرات. وتوفر حسابات تحليل المسار معلومات عن المسار الذي يأخذه متغير ما عند مساهمته السببية في متغيرات أخرى، كما توفر معلومات عن أهمية هذه المساهمة.

نظرية وطرق الاختبار التقليدية

يشترط في الاختبارات النفسية احتواؤها على عدد من الصفات، إلا أن أهم هذه الصفات هي صفة الصدق، إذ يمكن للاختبار أن يكون على درجة عالية من الكفاءة والثبات في القياس إلا أن هذا لا يعني شيئاً في حال فشل الاختبار في قياس ما قد صُمم لقياسه. وعندما يتم تصميم اختبار ما لتحقيق عدة أهداف يمكن أن يكون الاختبار صادقاً في قياس بعض هذه الأهداف ولكن ليس لأهداف أخرى. كما من الممكن أن يكون الاختبار مؤشراً دقيقاً وحساساً عن قدرة إدراكية معينة تحت ظروف معينة دون أخرى، أوأن يكون مقياساً أو أداة تنبؤ صادقة مع عينات معينة من الناس دون أخرى.

أنواع الصدق في الاختبارات:

يمكن الحصول على دلائل على صدق اختبار ما بالطرق التالية:

١- بواسطة تحليل محتوى الاختبار (صدق المحتوى).

٢- عن طريق دراسات تجريبية على درجة الإرتباط بين درجات الاختبار والدرجات على مقياس للمعيار تم الحصول عليه في نفس الوقت (الصدق التلازمي) أو يمكن الحصول عليه في المستقبل (الصدق التنبؤي).

٣- عن طريق تحليل درجة التطابق بين قدرة الاختبارعلى التمييز بين الأشخاص مع غيره من المقاييس لذات المفهوم العقلي أو تطابقه مع المواقف التي يكون فيها هذا المفهوم فعالاً (صدق المفهوم). وتستخدم الطرق الثلاث المذكورة أعلاه في إثبات قدرة الاختبار على قياس أو التنبؤ بالقدرات الإدراكية والسلوك في مجالات التعليم والتوظيف والطب وغيرها.

ركزت عمليات إثبات صدق الاختبارات التحصيلية المعايَرة بشكل رئيس على صدق المحتوى، إذ يتطلب وضع اختبار تحصيلي جيد إعداداً مسبقاً لجدول بالمواصفات أو مخطط موجز عن المحتوى المراد اختباره. وتعد مقارنة المحتوى النهائي للاختبار مع هذا الجدول الخطوة الأولى في عملية إثبات صدق محتوى الاختبار. كما يمكن الإستعانة بخبراء ومتخصصين في المادة موضع الاختبار لإبداء الرأي في درجة تطابق محتوى الاختبار مع أهدافه. ويمكن لعملية إثبات صدق المحتوى أن تقتصر على إصدار أحكام ذاتية تقرر ما إذا كانت أسئلة الاختبار تمثل المادة التعليمية موضع الاختبار بشكل جيد.

وعلى الرغم من أهمية تحليل محتوى الاختبار في تقييم مختلف أنواع الأدوات النفسية، يعد صدق المحك (والذي يتضمن الصدق التلازمي والصدق التنبؤي) الصفة الأساسية للاختبارات الجيدة التي تقيس الاستعدادات العامة والخاصة. ويجب تقديم دلائل على الصدق التنبؤي لاختبار ما قبل السماح باستخدام هذا الاختبار كأداة لاختيار أو تنسيب الموظفين والطلاب. وفي حال أثبتت الدراسات أن الاختبار تمكن من المساهمة بإزدياد فعالية التنبؤ بمحك الأداء المقبول بشكل اقتصادي، أي أن الاختبار يتميز بالصدق المتزايد، يتم تحديد معدل الدرجات الفاصلة للاختبار ولا يتم قبول أي ممتحن لايحقق هذا المعدل.

تستخدم أساليب الارتباط والإنحدار بشكل واسع في تحليل الصدق التنبؤي للاختبار، إذ غالباً ما ترتبط درجات الاختبارات التنبؤية بدرجات مقياس المحك أو تُستخلص منها. ويؤدي ارتفاع معامل الإرتباط بين أداة التنبؤ ومتغيرات المحك إلى ارتفاع دقة التصورالمستقبلي لدرجات المحك بالإستناد إلى درجات التنبؤ. ويمكن حساب، بالإضافة إلى الإرتباط بين الاختبار ومقياس المحك (r)، معادلة الإنحدار الخطي للتنبؤ بدرجات المحك بناءً على درجات الاختبار والخطأ المعياري للتقدير. ويمكن حساب الخطأ المعياري للتقدير (Sset) عن طريق المعادلة التالية:

$$s_{est} = s_y \sqrt{1 - r^2} \qquad (٢,١)$$

حيث Sy هي الإنحراف المعياري لمتغير المحك. ويمكن استخدام قيمة Sset في حساب مجال الخطأ في الدرجات التنبؤية للمحك والتي يمكن، وبدرجة معينة من الثقة، القول بأن الدرجات الناتجة للمحك تقع ضمنها حتماً.

ونظراً لأن الاختلاف بين درجات المحك التنبؤية والفعلية ممكن جداً فمن الجائز أن تقع أحدى الدرجات تحت حد الدرجات الفاصلة بينما تقع الثانية فوقه، وفي كلتا الحالتين يكون هناك خطأ في التنبؤ. ففي حال وقوع الدرجات التنبؤية للمحك فوق حد الدرجات الفاصلة ووقوع الدرجات الفعلية تحته ينتج ما يسمى بالخطأ الإيجابي غير الحقيقي، وفي حال وقوع الدرجات التنبؤية للمحك تحت حد الدرجات الفاصلة ووقوع الدرجات الفعلية فوقه ينتج ما يسمى بالخطأ السلبي غير الحقيقي. ومما لا شك فيه أن نقطة الحد الفاصل لمحك ما تؤثر على عدد الأخطاء الإيجابية والأخطاء السلبية الناتجة، إذ أن تحديد نقطة متدنية للحد الفاصل سوف يؤدي إلى عدد أكبر من الأخطاء الإيجابية غير الحقيقية بينما يؤدي تحديد نقطة مرتفعة للحد الفاصل للدرجات إلى عدد أكبر من الأخطاء السلبية غير الحقيقية. وتلعب عوامل أخرى دوراً في تحديد

الأخطاء السلبية والإيجابية غير الحقيقية. وتتمثل هذه العوامل في كل من: صدق الاختبار ونسبة المتقدمين الذين يجب اختيارهم (نسبة الاختيار)، والنسبة الدنيا (وهي معدل عدد المقبولين من مجمل تعداد العينة).

تساعد برامج الكمبيوتر A-3 وA-4 وA-5 الموجودة على القرص المرافق لهذا الكتاب في عملية تحليل الصدق التنبؤي للاختبارات إذ تحتوي هذه البرامج على إحصائيات للإرتباط والإنحدار تساعد على التحليل الإنحداري الخطي البسيط والمتعدد. وبالإضافة إلى تحديد معاملات الإرتباط ومعادلة الإنحدار الخطي، تساعد برامج A-4 وA-5 في حساب الخطأ المعياري للتقدير والاختبارات الإحصائية المناسبة لتحديد أهمية القيم المرجحة للإنحدار. ويحسب برنامج A-4 عدد الأخطاء السلبية والإيجابية غير الحقيقية المرتبطة بحد الدرجات الفاصلة.

يعد صدق المفهوم Canstruct Validity من أهم أنواع الصدق المرتبطة باختبارات الشخصية وغيرها من الاختبارات. ويعرّف صدق المفهوم على أنه قدرة الاختبار أو أداة التقييم على قياس سمة أوخاصية معينة تعتبر أساساً لتحديد السلوك البشري في مواقف معينة. فعلى سبيل المثال، تميز اختبارات القدرة اللغوية بين الأفراد الذين يمكن اعتبارهم مؤلفين ناجحين أو غير ناجحين بناءً على الاعتقاد بأن المؤلف الناجح يتمتع بقدرات لغوية أكثر تطوراً من المؤلف غير الناجح. إلا أن اثبات صدق المفهوم هو عملية أكثر تعقيداً من ذلك إذ أنها تتطلب عدداً من المواقف وعدداً من الدلائل إذ يعتمد صدق المفهوم على نتائج مجموعة من التحريات هدفها اثبات أن الأشخاص الذين حصلوا على درجات عالية في الاختبار تتوافق تصرفاتهم في عدد من المواقف مع ما هو متوقع من الأفراد الذين "يمتلكون" كمية كبيرة من هذا المفهوم. وفي الوقت ذاته تتوافق تصرفات أولئك الذين حصلوا على معدلات درجات متدنية في الاختبار مع ما هو متوقع من أولئك الذين يمتلكون كمية ضئيلة من المفهوم. ويشير صدق المفهوم إلى أن الارتباط بين

مقاييس المفهوم ذاته التي تم الحصول عليها بطرق مختلفة (المصداقية التقاربية) أعلى من الإرتباط بين عدة مفاهيم تم قياسها بالطريقة ذاتها (المصداقية التمييزية).

الثبات Reliability:

لا يمكن لاختبار أن يكون صادقاً إذا لم يقس المتغير ذاته في كل مرة.لا يعد ثبات الاختبار (أو توازن وثبات القياس)، شأنه كشأن الصدق، خاصية إما أن تكون موجودة في كافة أقسام الاختبار أو أن لا يكون لها وجود على الإطلاق، إذ يمكن للاختبار أن يكون أداة قياس ثابتة تحت ظروف معينة دون أخرى. وفي حال عدم كون الاختبار مقياساً ثابتاً لما يراد قياسه في الوقت الذي يجب أن يتم فيه القياس، لا يمكن اعتبار هذا المقياس صادقاً للخاصية أو الصفة أو السلوك موضع القياس. الثبات، على أية حال، هو شرط هام ولكن غير كافٍ لاثبات صدق الاختبار.

يرتبط ثبات الاختبار بكون معدلات الدرجات الناتجة عن هذا الاختبار معدلات ثابتة نوعاً ما على مر الزمن وتحت ظروف قياس متنوعة، وهذا لا يعني أن نتائج الاختبار لا تختلف بازدياد علم أو نضج أو مهارة الأفراد. وتجدر الإشارة إلى أن التركيز في عملية إثبات ثبات الاختبار يقع على الثبات ضمن مرحلة زمنية معينة في حين أنه في حالة الدراسات التطورية ينصب الاهتمام على مراقبة تطور الأفراد مع مرور الزمن. ويفترض في هذه الحالة أن نتائج الفرد بين فترة اختبار وأخرى تبقى ثابتة إلى حد ما وأنها خالية نسبياً من أخطاء القياس.

تفترض نظرية الاختبار التقليدية أن الدرجات الملاحظة للفرد (O) هي مجموع الدرجات الفعلية (T) وخطأ القياس (E)[٣] كما تفترض كذلك عدم وجود

(٣) تفترض نظرية الثبات التقليدية أن العلاقة بين الدرجات الفعلية وأخطاء القياس هي علاقة بين شيئين منفصلين يمكن إضافة الواحد منهما للآخر. وهذا في واقع الأمر قد =

ارتباط بين الدرجات الفعلية والأخطاء مما يعني أن تباين الدرجات الملاحظة يساوي مجموع تباين الدرجات الفعلية (الحقيقية) وتباين خطأ القياس:

$$s_o^2 = s_t^2 + s_e^2 \qquad (٢-٢)$$

تعرّف الثقة (r_{11}) بناءً على النموذج أعلاه على أنها نسبة تباين الدرجات الفعلية إلى تباين الدرجات الملاحظة:

$$r_{11} = s_t^2 / s_o^2 \qquad (٢-٣)$$

ومن الجدير بالذكر أن معظم المعادلات في نظرية الثبات التقليدية تستند إلى المعادلتين أعلاه.

لا يمكن حساب معامل الثقة بشكل مباشرة باستخدام المعادلة ٢-٣ بسبب جهل التباين الفعلي لدرجات الاختبار، إلا أن عدداً من طرق الحساب غير المباشرة قد صممت لتقدير ثبات الاختبارات. ويمكن تحديد أربعة أنواع من تقديرات الثبات هي: معاملات الثبات، والتكافؤ، والثبات والتكافؤ، والاتساق الداخلي. يُحسب معامل الثبات أو معامل الاختبار – إعادة الاختبار عن طريق إجراء الاختبار لذات المجموعة من الأشخاص مرتين مختلفتين ثم إيجاد الإرتباط بين الدرجات في المرتين. أما معامل التكافؤ أو معامل الصور المتكافئة فيتم

= يكون افتراضاً غير صحيح إذ يمكن للعلاقة بين الدرجات الفعلية وأخطاء القياس أن تكون علاقة بين شيئين تابعين أو مترابطين يمكن مضاعفتهما. فعلى سبيل المثال، يمكن للدرجات الملاحظة لعدد من الأفراد يتمتعون بقدرة إدراكية عالية أن تحتوي على نسبة خطأ أعلى من معدلات الأفراد ذوي القدرات الإدراكية الضعيفة. وتساهم عوامل كالدافع الداخلي أو الرغبة في أخذ الاختبار والسرعة وتوزيع الوقت وغيرها من العوامل اللاإدراكية، التي تتنوع بتنوع مستويات أداء الأفراد في أخطاء القياس الإدراكي. وتعطي الافتراضات التجميعية والمستقلة لنظرية الثبات التقليدية أبسط تصور لمفهوم نموذج العلاقة بين الدرجات الفعلية وأخطاء القياس، وبالرغم من توافق هذا التصور مع النموذج الخطي العام في نظرية الإحصاء إلا أنه ليس بالضرورة أن يكون التصور الأكثر صحة.

حسابه عن طريق تصميم صورتين متشابهتين من الاختبار واستخدامهما لاختبار المجموعة نفسها من الممتحنين ثم حساب الارتباط بين درجاتهما. كما يتم حساب معامل الثبات والتكافؤ عن طريق إجراء صورتين متشابهتين من الاختبار للمجموعة ذاتها من الأفراد في توقيتين مختلفين. وللتحكم في تأثيرات نظام إدارة الاختبار يتم تقسيم المتقدمين للاختبار إلى مجموعتين: المجموعة الأولى تأخذ الاختبار (أ) في الوقت (١) والاختبار (ب) في الوقت (٢)؛ أما المجموعة الثانية فتعطى الاختبار (ب) في الوقت (١) والاختبار (أ) في الوقت (٢). ويتم بعد ذلك حساب الإرتباط بين درجات نموذجي الاختبار للمجموعتين.

ولا تخلو الطرق الثلاثة المذكورة أعلاه من العلل، إذ يمكن لطريقة الاختبار وإعادة الاختبار أن تعطي بعض الممتحنين (وليس جميعهم) خبرة في الاختبار مما قد يؤدي إلى تحسن درجاتهم في المرة الثانية مما يؤثرعلى حساب الإرتباط بين الدرجات في كلتا المرتين. وأما في الطريقتين الثانية والثالثة فيعتبر إعداد نموذج مشابه أو صورة مكافئة للاختبار عملية معقدة وشاقة ولا تضمن التكافؤ الحقيقي بين نموذجي الاختبار. ولهذه الأسباب تم بذل جهود منذ بدايات عملية القياس النفسي إلى ترشيد عملية تقدير ثبات الاختبارات. ومن ضمن هذه الجهود محاولة إيجاد طرق لتحديد الاتساق الداخلي لاختبارٍ ما – أو درجة قياس العناصرالتي يتألف منها الاختبار للمتغير ذاته.

يفترض في الاختبارات التي تحتوي على عدد كبير من الفقرات التي تقيس الشيء ذاته أن تكون على درجة أعلى من الثبات بسبب كون الاختبارات ذات عدد الفقرات القليل أكثر عرضة لأخطاء القياس من الاختبارات الطويلة. وبالإضافة إلى ذلك يمكن تقدير الثبات في النماذج المتكافئة من الاختبار عن طريق وضع درجات الاختبار على مرحلتين، المرحلة الأولى للأسئلة ذات الأرقام الزوجية والمرحلة الثانية للأسئلة ذات الأرقام الفردية. يتم بعد ذلك حساب الإرتباط بين

درجات نصفي الاختبار $r_{1/2 1/2}$ ثم تصحح على أساس طول الاختبار الكامل (وهو ضعف طول كل نصف) وفق معادلة سبيرمان – براون:

$$r_{1/2 1/2} = 2r_{1/2 1/2} / (1 + r_{1/2 1/2}) \qquad (٢-٤)$$

ويعد حساب معامل ثبات التجزئة النصفية الذي يتم الحصول عليه وفق المعادلة أعلاه مقياساً للثبات الداخلي للاختبار أو لدرجة قياس فقرات الاختبار للمتغير ذاته.

طور كودر وريتشاردسون إجراءاً أكثر شمولية لحساب معاملات الاتساق الداخلي لفقرات الاختبار ثنائية الدرجة التي تحتمل الدرجات ١-٠ (١ في حال صحة الإجابة و٠ في حال خطئها)، ويطلق على هذا الإجراء اسم معادلات كودر- ريتشاردسون. وقد عمم كرونباخ إجراءات كودر- ريتشاردسون لتشمل الفقرات الاختبارية ذات الدرجات المتعددة (أكثر من ١ أو صفر) مما أنتج المعامل ألفا:

$$a = k(1 - \sum s_i^2 / s_o^2)/(k-1) \qquad (٢-٥)$$

حيث k هي عدد الفقرات وs_i^2 هي تباين الفقرة i وs_o^2 هي تباين مجموع درجات الاختبار؛ والمجموع ($\sum$) هو لجميع الفقرات k.

ويمكن حساب معاملات سبيرمان -براون وكودر- ريتشاردسون ومعامل كرونباخ ألفا بواسطة برامج الكمبيوتر E-1 و E-2 و E-3 الموجودة على القرص المرافق لهذا الكتاب. كما يساعد برنامج E-6 في حساب الخطأ المعياري للقياس و نسب ٩٠٪ و٩٥٪ و٩٨٪ ٩٩٪ لمجال ثقة الدرجات الفعلية الموافقة للدرجات الملاحظة المعطاة لاختبار ما. ويمكن الحصول على معادلة الخطأ المعياري للقياس (وهي إجراء إحصائي عام يطبق على درجات جميع الممتحنين) بالطريقة

التالية: إن حل s_t^2 في المعادلة (٢-٥) يعطينا $s_t^2 = r_{11}s_o^2$. وبتعويض هذا الناتج بـ s_t^2 في المعادلة (٢-٢) وحل s_e^2 نحصل على: $s_e^2 = s_e^2 - r_{11}s_o^2$.

وأخيراً، يعطينا تبسيط طرفي المعادلة وأخذ الجذر التربيعي للطرفين الخطأ المعياري للقياس:

$$s_e = s_o\sqrt{1 - r_{11'}} \qquad (٢-٦)$$

ويطبق هذا الإجراء الإحصائي بالطريقة التالية: بفرض أن ثبات الاختبار يساوي٩٠٪ وأن الإنحراف المعياري هو ١٠، فإن الخطأ المعياري للقياس هو:

$10\sqrt{1-.9} = 3.16$. ولحساب المجال الذي تقع ضمنه، وبمقدار ٩٥٪من الثقة، الدرجات الفعلية لفرد ما هو عضو في مجموعة درجاتها الملاحظة، نقوم بالعملية الحسابية التالية: $O \pm 1.96(s_e) = 43.81, 56.19$.

إن مفهوم الثبات ليس محدوداً بالمتغيرات الفردية إذ يمكن تحديد مدى ثبات التباين بين درجات متغيرين أو ثبات مركّب يتألف من متغيرين أو اكثر. ويمكن حساب الثبات والخطأ المعياري لقياس درجات التباين عن طريق برنامج الكمبيوتر E-7. تساعد هذه الإحصائيات في تحديد ما إذا كانت درجات فردٍ ما في اختبارين مختلفين أو في اختبارين فرعيين من نفس الاختبار (كبطارية الاستعدادات المتعددة) تختلف عن بعضها بشكل كبير. وتؤكد النتائج على أن ثبات التباين بين درجات اختبارين هي أقل من ثبات كل اختبار على حدة، وعليه فإن الخطأ المعياري لقياس تباين الدرجات هو أعلى بكثير من الخطأ المعياري للقياس في كل من الاختبارين.

يمكن استخدام برنامج E-8 في تقديرثبات اختبار مؤلف من عدة أقسام لكل قسم منها ثباته وإنحرافه المعياري وثقله العددي. ويمكن للبرنامج حساب مجموعة الأثقال المثلى التي تعطي أقصى درجة من الثبات المركّب.

نظرية التعميم

يساعد التحليل العاملي على تحليل تباين الدرجات الفعلية في المعادلة (٢-٢) إلى تباينات عاملية عامة وتباينات عاملية خاصة، إلا أن هذه المعادلة لا تميز بين مصادر الخطأ المختلفة التي تساهم في تباين الخطأ (Se^2). وقد أثبت خبراء القياس النفسي على مر السنين أن للاختبارات أنواع متعددة من الثبات تعتمد على مصادر الخطأ في القياس التي تُأخذ بعين الاعتبار عند حساب معاملات الثبات.

تؤثر العوامل التالية على تقدير الثبات: الفقرات المحددة التي يتكون منها الاختبار وتعليمات الاختبار والظروف المحيطة بالممتحنين ومواصفات عينة الاختبار. كما تعطينا الطرق المختلفة في حساب الثبات (طريقة الاختبار وإعادة الاختبار والنماذج المكافئة والاتساق الداخلي) نتائج مختلفة.

وقد اعتمدت الإتجاهات الحديثة لحساب الثبات على الإنطلاق من حقيقة أن عدداً من مصادر الخطأ المختلفة (المناسبات أو النماذج، أو المصححون، أو الإداريون) يمكن أن تؤثر على ثبات الاختبار، إذ يمكن للاختبار أن يكون ثابتاً في عدد من النواحي حسب تأثير هذه المصادر أو المظاهر. وتفيد نظرية التعميم، وهي نظرية جديدة لحساب صدق وثبات المقاييس الإدراكية، بأن الدرجات الملاحظة ما هي إلا تقدير لمعدل درجات عام (Universe Score). وإن مدى صحة درجات اختبار ما في تقدير قيمة عامة ما هي إلاّ مقياس لثبات الاختبار. ويعتمد مدى إمكانية التعميم من درجات الاختبار الملاحظة إلى القيمة العامة على طبيعة الوسط المراد التعميم عليه أي على المظاهر التي تعرّف هذا الوسط.

يمكن تعميم درجات اختبار ما على عدد من الأوساط العامة يعرّف كل واحد منها تبعاً للمظاهر التي تدخل في تركيبه، وقد تختلف جذرياً من وسط لآخر كل

من العناصر التي تشخِّص وسطاً عاماً معيناً مثل ظروف إدارة الاختبار، وتركيبة الفقرات أو نماذج الاختبار أو ما شابه ذلك. وفيما لا تؤثر بعض هذه المظاهر على تعميم درجات الاختبار، يلعب البعض الآخر دوراً شديد الأهمية في عملية التعميم. كما يمكن الحصول على معاملات التعميم أوالثبات عند تعميم الدرجات على أوساط متعددة المظاهر.

وتتطلب عمليات الحساب الإحصائية للتعميم تطبيق تحليل لطرق تباين مصممة لتحديد دلالة عناصر معينة وتأثيرها على معاملات التعميم الناتجة. وبتأكيدها على أهمية الظروف المحيطة بالاختبار من حيث طريقة إجراء الاختبار والأهداف التي صمم من أجلها، ساهمت نظرية التعميم في إزاحة الهاجس التقليدي لمستخدمي الاختبار الذي كان الشغل الشاغل للعاملين في مجال الاختبارات والمتعلق بكون الاختبار اختباراً جيداً أو سيئاً بهاجسٍ جديد وهو: "ما هو الغرض الذي يعتبر هذا الاختبار جيد أو سيئ له؟".

قياس الاختبارات والمعايير

إن درجات الاختبارات وغيرها من المقاييس النفسية لا تعني شيئاً بمفردها، إذ يتوجب تفسير هذه الدرجات وفق المقاييس الذاتية الخاصة بالأداء الجيد أوالسيئ أو الطبيعي أو غير الطبيعي التي يمتلكها أفراد مؤهلون لإصدار مثل هذه الأحكام. ويتم التوصل إلى هذه المقاييس الذاتية عن طريق الخبرة في مادة الاختبار وعن طريق توقع ما يمكن اعتباره مستوى الأداء المقبول. فعلى سبيل المثال، يمتلك المدرسون ذوي الخبرة الخاصة بالمادة والطلاب، وعليه فقد أنشؤوا إطاراً داخلياً مرجعياً والذي بواسطته أصبح بإمكانهم مقارنة الإجابات المختلفة التي يعطيها الطلاب ودرجات الاختبار لمجموعة طلاب. إلا أن هؤلاء المدرسون الخبراء يختلفون مع الأسف فيما بينهم بخصوص تفسير الدرجات ووضع الدرجات النهائية ووضع درجات الاختبارات المقالية وغيرها من مقاييس الأداء ذات الطبيعة غير الموضوعية.

اتضحت أهمية إيجاد طرق موضوعية لتقييم درجات الاختبارات منذ بدايات تاريخ الاختبارات النفسية والتعليمية. ونظراً لعدم صلاحية الاعتماد على خبرة المصحح التي لا تقترن بالوثائق في تقييم الأداء الاختباري لمجموعة كبيرة من الأفراد، تم اعتماد عملية إجراء الاختبارات لعينة كبيرة من الأشخاص تمثل المجموعة السكانية المراد اختبارها. ويتم انتقاء أفراد العينة عشوائياً بشكل يتناسب مع التمثيل الديموغرافي للأفراد من الأعمار والصفوف الدراسية والأعراق المختلفة المكونة للمجتمع (وذلك لأن هذه المتغيرات تؤثر على درجات الاختبارات). ويساعد برنامج D-1 الموجود على القرص المرافق لهذا الكتاب في اختيار عينات عشوائية طبقية أو متعددة الطبقات والذين سيستخدمون كأداة لقياس الاختبار، ويتم إجراء الاختبارات القياسية تحت ظروف موحدة ومُسيطر عليها وذلك لأن شخصية وأسلوب القائمين على الاختبارات والظروف المحيطة بالاختبار قد يكون لها آثار واضحة على نتائج الاختبارات، وتتم عملية التصحيح بعد الفراغ من إجراء الاختبار ثم يتم إعداد التوزيع التكراري للدرجات. بعد ذلك، يتم تحويل الدرجات إلى أنواع متعددة من المعايير تستخدم لاحقاً كمقياس تُقارن معه درجات الذين سيتم اختبارهم في المستقبل. ومن الضروري التأكد من أن عملية تقرير درجات ممتحن ما تستند إلى مقارنة درجاته مع جدول المعايير الناتج عن إجراء الاختبار لعينة من الأفراد تشابه خصائص أفرادها خصائص هذا الممتحن، وذلك بالنظر إلى الخصائص السكانية. فعلى سبيل المثال يقارن الأداء الاختباري لطالب من الصف الخامس مع المعايير التي تم الحصول عليها من عينة من طلاب الصف الخامس وليس الصف الرابع أو السابع، ويقارن الأداء الاختباري لطفل في العاشرة من عمره مع أداء أطفال في العاشرة أو مع أداء طلاب في الصف الخامس.

وتعد معايير العمر والسنة الدراسية (الصف الدراسي) والمعايير المئينية ومعايير الدرجات القياسية من أهم المعايير التي تحتويها دلائل المستخدم

للاختبارات المختلفة. ويتم حساب معيار العمر عن طريق حساب وسيط درجات لأطفال من العمر الزمني المعطى. فإذا كان وسيط الدرجات لأطفال عمر كل منهم ١٠ سنوات و٥ أشهر هو ٦٠، فإن معيار العمر لفرد يحصل على معدل الدرجات ٦٠ في هذا الاختبار هو ١٠ سنوات و٥ أشهر. وتحتوي اختبارات الذكاء القديمة على نوع آخر من المعايير العمرية هو معيار العمر العقلي.

وكما هو الحال في المعيار العمري يتم حساب معيار الصف الدراسي عن طريق حساب وسيط درجات طلاب سنة دراسية معينة. فإذا كان الوسيط لدرجات طلاب في الشهر السادس من الصف الرابع يساوي ٥٥، فإن معيار الصف لفرد يحصل على معدل الدرجات ٥٥ في هذا الاختبار هو ٤,٦، ومن الجدير بالذكر أن التعبير عن معيار الصف الدراسي هو بالسنوات الدراسية وأجزاء من العشرة، مما يعني أفتراض عدم تطور الأطفال في أشهر العطلة الصيفية.

وعلى الرغم من شيوع استخدام طريقتي معيار العمر ومعيار الصف في تفسير درجات الاختبارات القياسية، يبقى هنالك خطر التعامل مع هذا المستوى الرتبي للقياس على أنه مستوى فئوي. إذ أن القول بأن الفرق بين مقياس صف ٤,٦ و مقياس آخر يساوي ٦,٦ هو مساوٍ للفرق بين المقياسين ٦,٦ و٨,٦ هو استنتاج خاطئ لسبب بسيط وهو أن القدرات الأكاديمية تتطور بشكل أكبر وأسرع في السنوات الأولى من عمر الطالب وفي سنوات الدراسة الأولى بالمقارنة مع السنوات اللاحقة. وعليه لا تتساوى قيم وحدات معيار العمر ومعيار الفصل الدراسي في مختلف سنوات العمر والمراحل الدراسية.

تتصف المعايير المئينية (وهي نوع آخر من الدرجات النهائية للاختبارات) بعدم وجود علاقة خطية (منتظمة) بين الدرجات فيها والزمن. والمعيار المئيني أو معيار الرتب المئيني لاختبار ما هو النسبة المئوية لعدد أفراد عينة القياس الذين حصلوا على معدل الدرجات هذا أو على معدل أدنى منه. وعليه فإذا حصل ٦٠٪ من أفراد عينة القياس على معدل درجات ٧٥ أو أقل فإن الرتبة المئينية لفرد

معدل درجاته ٧٥ هو ٦٠، ويتوجب عند تحويل معدلات الدرجات إلى رتب مئينية مراعاة تناسب عينة القياس التي تم تحديد المعايير على أساسها مع الممتحن المراد تفسير درجاته.

ولعل أحد أسباب شيوع استخدام معايير العمر والصف والمعايير المئينية كون هذه المعايير سهلة الفهم لغير المختصين في مجال القياس النفسي. أما قواعد الدرجات المعيارية فهي مقاييس أكثر دقة من وجهة النظر السايكومترية إلا أنها أكثر تعقيداً. وتختلف قواعد الدرجات المعيارية عن قواعد العمر والصف والمعايير المئينية بكونها مقاييس فئوية، إلا أنها لم ترقى إلى المرتبة النِسَبية لقياس. تستند جميع الدرجات المعيارية إلى معدل درجات z على الشكل التالي:

$$z = (X - \overline{X})/s \qquad (٧-٢)$$

حيث X هي معدل الدرجات الخام و$\overline{X}$ هي الوسط الحسابي وs هي الإنحراف المعياري. وبما أن معدلات الدرجات z يمكن أن تكون أعداداً عشرية إما موجبة أو سالبة يتم ضربها بعدد صحيح ثم يضاف عدد صحيح آخر إلى الناتج. فعلى سبيل المثال، يتم حساب معدلات الدرجات T بضرب معدلات الدرجات z ب ١٠ ثم إضافة ٥٠ للناتج. ويتم حساب انحراف معدلات ذكاء ويكسلر بضرب معدلات الدرجات z ب ١٥ ثم إضافة ١٠٠، وتحسب معدلات درجات مجلس امتحانات القبول الجامعية (CEEB) بضرب معدلات الدرجات z ب ١٠٠ ثم إضافة ٥٠٠، وتحسب درجات الرتب من تسعة (Stanine) بضرب معدلات الدرجات z ب ٢ ثم إضافة ٥. وهنالك عدة أنواع من الدرجات المعيارية مثل NCE (الدرجات المكافئة للمنحنى الطبيعي) والدلتا. تفترض العديد من هذه الدرجات المشتقة أن التوزيع التكراري للدرجات هو توزيع طبيعي، ولذلك فإنه من المناسب استخدام الدرجات القياسية المطبّعة Normalized للتعبير عن هذه الدرجات.

تقيس البرامج ٢ و٣ و٤ من الفئة D الموجودة على القرص المرفق بهذا الكتاب المعايير المئينية ومعايير الدرجات القياسية ومعايير الدرجات القياسية المطبّعة. ويعد برنامجا D-2 وD-4 برنامجين متعددي المزايا إذ يمكن استخدامهما في اتجاهين، كاستخدام برنامج D-2 لتحديد الرتبة المئينية لمعدل درجات ما أو معدل الدرجات الموافق لرتبة مئينية ما. ويحدد البرنامج D-4 إما الاحتمالية الطبيعية لقيمة z أو القيم z و T و CEEB وستانين ومعدلات درجات NCEالموافقة لاحتمالية طبيعية تراكمية.

تحليل فقرات الاختبار

نظراً لأن درجات الاختبارات التحصيلية واختبارات الاستعداد المختلفة غالباً ما تعتمد على حساب مجموع أسئلة أو فقرات الاختبار التي أجاب عليها المتحن بشكل صحيح، فإن صدق وثبات مجموع درجات الاختبارات تعتمد على دلالات الفقرات أو العناصر المكونة لهذه الاختبارات. وفي حالة الاختبارات التحصيلية حيث يتم التركيز على صدق المحتوى وعلى الإتساق الداخلي للاختبار، فإن تقييم دلالات الفقرات الاختبارية يتم عن طريق تحديد الإرتباط بين أداء الفرد في هذه الفقرة وأداؤه في الاختبار بشكل عام. ويمكن الحصول على معلومات تحليل فقرات الاختبار عن طريق استخدام مؤشر صعوبة الفقرات ومؤشر تمييز الفقرات. ويتم حساب مؤشر صعوبة الفقرات (p)، والذي يعبرعن نسبة الأفراد ضمن عينة الاختبار الذين أعطوا إجابة صحيحة على سؤال ما، عن طريق المعادلة $p = n/N$ حيث n هي عدد الأفراد الذين أعطوا إجابة صحيحة وN هي عدد الأفراد الذين تم اختبارهم. ويمكن حساب مؤشر الصعوبة بطريقة أخرى وهي تقسيم مجموعة المتحنين إلى مجموعتين فرعيتين هما مجموعة تضم أفراداً معدلات درجاتهم تضع في أفضل ٢٧٪ من المعدلات (U) ومجموعة تضم أفراد معدلات درجاتهم تقع في أسوأ ٢٧٪ من المعدلات (L)

وذلك تبعاً للتوزيع التكراري التالي لمجموع معدلات الدرجات. ويتم حساب p كما يلي:

$$p = (U_c + L_c)/(U + L) \qquad (٢-٨)$$

حيث Uc هو عدد الأفراد في المجموعة الأولى الذين أعطوا إجابة صحيحة على السؤال وLc هو عدد الأفراد من المجموعة الثانية الذين أجابوا على السؤال بشكل صحيح. وبالإضافة إلى مؤشر صعوبة فقرات الاختبار، يتم حساب مؤشر التمييز للفقرات الاختبارية (D) وهو إجراء للتمييز بين المجموعتين الأولى والثانية من حيث إجابات أفرادهم على فقرة امتحانية ما - كما يلي:

$$p = (U_p + L_p)/(U + L) \qquad (٢-٩)$$

وبشكل عام، تكون مؤشرات الفقرات الاختبارية الفعالة كالتالي: يتراوح مؤشر p بين ٢٠ و٨٠ وتكون مؤشرات D٣٠ وما فوق. يساعد برنامجC-I الموجود على القرص المرفق بهذا الكتاب في حساب مؤشرات الصعوبة والتمييز لأعداد كبيرة من الفقرات الامتحانية.

تقوم الاختبارات ذات المرجعية المحكِّية Criterion - referenced tests، وعلى عكس الاختبارات ذات المرجعية المعيارية Norm - referenced tests، على أساس مقارنة الإجابات مع مجموعة من الأهداف التعليمية المراد تحقيقها وليس على أساس مقارنة الإجابات بعضها مع بعضها. وبينما يتم حساب مؤشر الصعوبة لفقرات الاختبارات ذات المرجعية المحكِّية بنفس الطريقة المستخدمة مع الأنواع الأخرى من الاختبارات، يحسب مؤشر التمييز وفق المعادلة التالية:

$$D = U_c / U - L_c / L \qquad (٢-١٠)$$

حيث U هي عدد الممتحنين الذين حصلوا على معدلات درجات أعلى من مستوى المحك، وL هي عدد الممتحنين الذين حصلوا على معدلات درجات أدنى من

مستوى المحك، وهو عدد الأفراد في المجموعة U الذين نجحوا في الفقرة الاختبارية، وL_c هو عدد الأفراد في المجموعة L الذين نجحوا في الفقرة الاختبارية. ويساعد برنامج C-2 الموجود على القرص المرفق بهذا الكتاب في حساب مؤشرات الصعوبة والتميز لفقرات للاختبارات ذات المرجعية المحكية.

ويستخدم المعامل الثنائي الأصيل Point - biserial coefficient (r_{pb}) بدلاً عن مؤشر تمييز الفقرات عندما يكون المحك الذي يقيِّم فعالية الفقرة الاختبارية محكاً خارجياً كمقاييس الأداء الوظيفي أو الأكاديمي. ويتم حساب هذا المؤشركالتالي:

$$r_{pb} = (\bar{Y}_c - \bar{Y})\sqrt{Nn_c/[(N-n_c)(N-1)]}/s. \qquad (٢-١١)$$

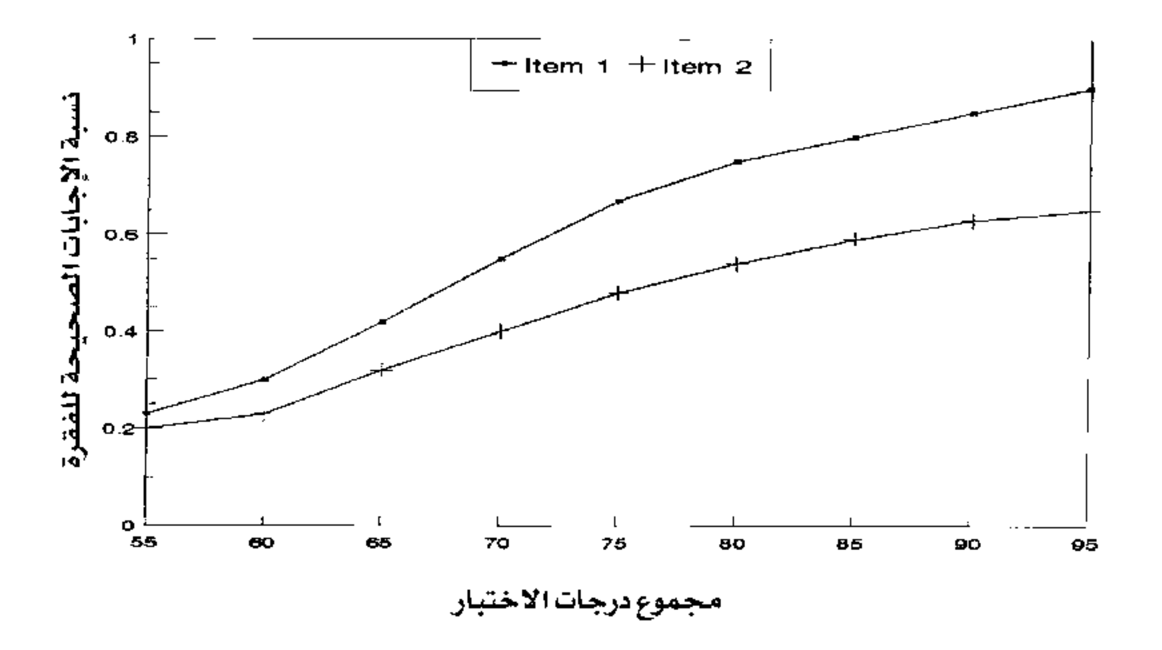
حيث N تساوي مجمل عدد المتحنين وn_c هي عدد المتحنين الذين أجابوا على السؤال بشكل صحيح و$\bar{Y}_c$ هي متوسط الدرجات على مقياس المحك للممتحنين الذين أجابوا على الفقرة الامتحانية بشكل صحيح و$\bar{Y}$ هي متوسط درجات المحك لجميع المتحنين و sهي الإنحراف المعياري لدرجات المحك لكافة المتحنين. وعلى الرغم من عدم وجود قيمة محددة لـ (r_{pb}) يتم على أساسها استبعاد الفقرات الاختبارية التي لا تصل إلى هذه القيمة، يمكن القول أن عدد نقاط متدنٍ جداً لـ (r_{pb}) يمكن أن يؤدي إلى مراجعة الفقرة الاختبارية أو الاستغناء عنها. ويساعد برنامج C-2 الموجود على القرص المرفق بهذا الكتاب في حساب المعامل الثنائي الأصيل.

يمكن الحصول على معلومات أكثر فائدة وتفصيلاً عن تحليل الفقرة الاختبارية وذلك عن طريق رسم منحنى لخصائص الفقرة (ICC) يعبر عن نسبة الإجابات الصحيحة إلى الفقرة معبراً عنها بمجموع درجات الاختبار (الشكل ٢-٢). نلاحظ في الشكل ٢-٢ توافُق قيمة ٠,٥ على المحور الرأسي مع مجموع

درجات مقداره ٦٨ لفقرة الاختبار ١ و٧٧ لفقرة الاختبار ٢، وهذا يعني أن الفقرة ٢ أكثر صعوبة من الفقرة ١، إلا أن منحنى الفقرة ١ يتصف بانحدار أشد من الفقرة ٢ مما يشير إلى أن الفقرة ١ أقدر على التمييز بين المتحنين. ويساعد برنامج C-4 الموجود على القرص المرفق بهذا الكتاب في حساب منحنيات ICC بناءاً على المعلومات الأولية المأخوذة من نموذج الإجابة على فقرات الاختبار ذات المتغيرات الأحادية، أو الثنائية أو الثلاثية (راجع الفصل التالي).

الشكل ٢-٢

منحنيان لخصائص الفقرات لهما مؤشران مختلفان للصعوبة والتمييز



وبالإضافة إلى حساب p وإما قيم D أو (r_{pb}) لكل فقرة من فقرات الاختبار، يجب النظر في التوزيع التكراري للمموهات itemdisracters (الخيارات غير الصحيحة) في أسئلة الاختيار من المتعدد ضمن المجموعتين U و L. تعد المموهات في أسئلة الاختيار المتعدد عنصراً مهماً من عناصر فعالية الفقرة الاختبارية، وقد

تؤثر هذه المموهات تأثيراً هداماً على الفقرة الاختبارية إذا أدت إلى الغموض أو كانت غير فعالة. إذ يمكن أن تقود الأخطاء أو التناقضات أو غيرها في محتوى المموهات إلى اختيار نسب كبيرة (وغير متوقعة) من الممتحنين لإجابة مموهة ما بدلاً عن الإجابات الأخرى مما يؤدي إلى سوء في توزيع اختيار المموهات وإلى إعادة النظر في الفقرة الامتحانية المعنية أو إلى حذف هذه الفقرة بشكل كامل. ويساعد برنامج C-3 الموجود على القرص المرفق بهذا الكتاب في إجراء اختبار إحصائي حول توزيع الإجابات على مموهات أسئلة الاختيار المتعدد.

وأخيراً؛ إن كون مؤشري الصعوبة والتمييز لاختبار ما مقبولين، لا يعني أن هذا الاختبار يعمل بالشكل نفسه مع كافة المجموعات البشرية (من عرق أو جنس أو عمر معين) التي يتم اختبارها. ولهذا السبب يتطلب وضع الاختبارات التحصيلية القياسية وغيرها من اختبارات القدرة الحصول على معلومات تحيز Biasedness فقرات الاختبار أو ما يسمى بفعل الفقرة التبايني (DIF) Differential item functioning. يحتوي برنامج C-7 على مقاربات للحصول على معلومات فعل الفقرة الاختبارية التبايني عن طريق: الرسم البياني لدلتا الفقرة وطريقة كاي تربيع التي استخدمها شيونمان وكاميللي ومانتل وهاينتسل. وتعد طريقة الرسم البياني لدلتا الفقرة الاختبارية أقدم هذه الطرق إلا أن طريقة كاي تربيع التي استخدمها مانتل وهاينتسل كانت الأكثر استخداماً خلال السنوات الأخيرة.

درج صانعوا الاختبارات على الاحتفاظ بنتائج تحليل الفقرات الاختبارية على بطاقات مفهرسة ترفق بالفقرة الاختبارية وتخزن بغرض استخدامها لاحقاً. أما في الوقت الحالي ومع وجود أجهزة كمبيوتر ذات سرعة المعالجة العالية فيتم تخزين الفقرات الاختبارية ضمن بنك تصنف فيه تبعاً للموضوع أو درجة الصعوبة أو مستوى التمييز أو بحسب العمليات الإدراكية المطلوبة للإجابة عليها. ولا يقتصر استخدام "بنوك" فقرات الاختبار هذه على واضعي الاختبارات

المحترفين إذ غالباً ما ترفق هذه "البنوك" بالعديد من الكتب الدراسية بغرض استخدامها كأسئلة تدريبية أو كمصدر لإعداد الاختبارات المدرسية. ويمكن لمدرسي المقررات أن يستخدموا حواسيبهم المكتبية لاختيار فقرات اختبارية متعلقة بعدد من المضامين ثم تجمع هذه الفقرات معاً لتشكل اختبارات وحدات دراسية معينة أو امتحانات عامة. وتفيد برامج الحاسب في تسهيل عملية إختيار فقرات اختبارية مرتبطة بمادة دراسية معينة وتمتلك عدداً من المواصفات الإحصائية المحددة.

نظرية الاستجابة لفقرة الاختبار IRT Item response theory

تختلف نظرية الإستجابة للفقرة عن غيرها من نظريات وطرق الاختبار التقليدية التي لا تعير اهتماماً كبيراً للإجابات على الفقرات الاختبارية، إذ أن التركيز الأساسي لنظرية IRT هو على منهجية الإجابة على فقرات الاختبار. وتستند هذه النظرية التي تعرف باسم IRT إلى العلاقة الوظيفية النظرية بين سلسلة للقدرات الكامنة عند الفرد وبين الإجابات على فقرات الاختبار. ويتم تمثيل نموذج IRT عن طريق تابع لوجستي ذو مؤشر واحد أو مؤشرين أو ثلاثة مؤشرات. ويعبر عن النموذج ذو المؤشرات الثلاثة كالتالي:

$$P(\theta) = c + (1-c)/[1+e^{-a(\theta-b)}] \qquad (١٢-٢)$$

حيث e هي قاعدة اللوغاريتمات الطبيعية (٢,٧١٨٢٨٢) وb هي مؤشر صعوبة الفقرة الاختبارية و c هي مؤشر التخمين الكاذب وθ هي مستوى قدرة المجيب على السؤال (الممتحن) بالنسبة لمقياس الدرجات المعيارية و$P(\theta)$ هي احتمالية إجابة فرد ما مستوى قدرته هو θ بشكل صحيح على السؤال. وبفرض أن $c=0$ نحصل على النموذج ثنائي المؤشرات:

$$P(\theta) = 1/[1+e^{-a(\theta-b)}] \qquad (١٣-٢)$$

ويؤدي الافتراض بأن جميع الفقرات الاختبارية لها درجة التمييز ذاتها إلى النموذج أحادي المؤشر أو ما يسمى نموذج راش:

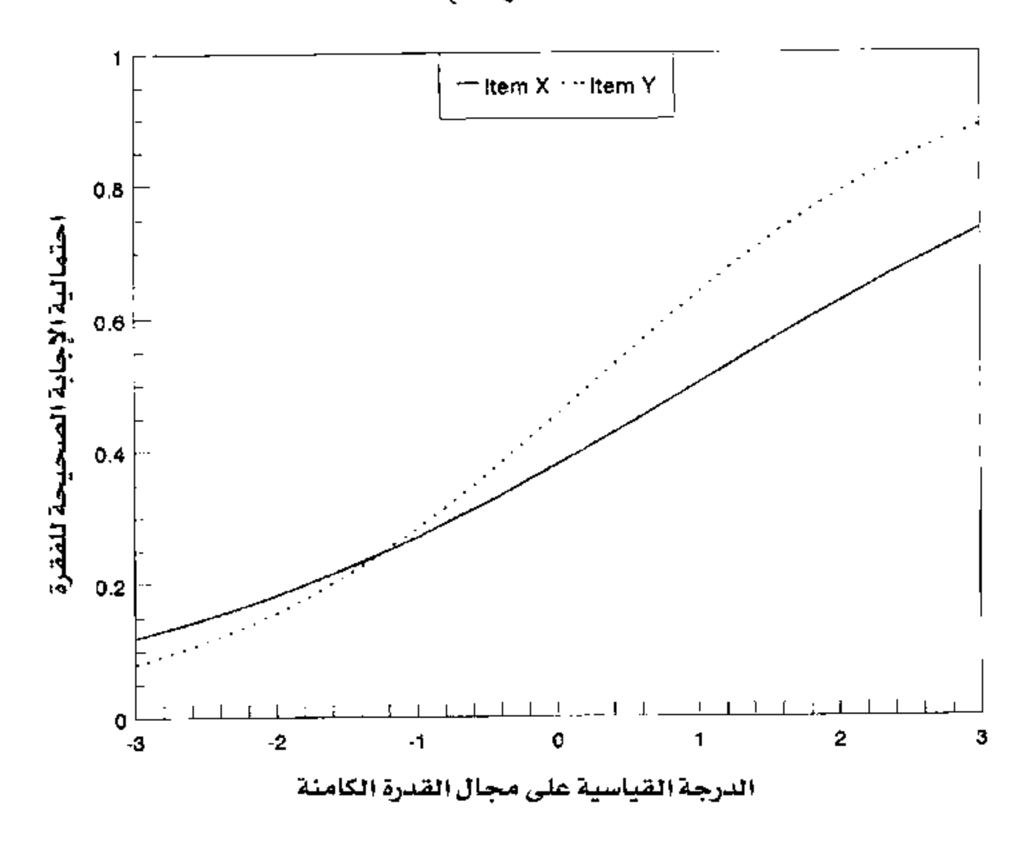
$$P(\theta) = 1/[1 + e^{-(\theta - b)}] \qquad (٢-١٤)$$

وعلى الرغم من أن نموذج راش كان أساساً للعديد من الأبحاث في مجال القياس النفسي، يبقى النموذج ثنائي المؤشر هو النموذج الأكثر شيوعاً.

وكما يوضح الشكل ٢-٣، يختلف شكل المنحنى الذي ترسمه الإجابات على فقرة اختبارية ما باختلاف قيم المؤشرات a وb. وتم رسم المنحنيين في الشكل ٢-٣ بالاعتماد على التابع ذو المؤشرين الموجود في المعادلة ٢-١٣ وكان مؤشرالصعوبة (b) هو ١,٠٠ ومؤشرالتمييز (a) هو ٠,٥٠، على المنحنى X، على حين أن قيم b و a على المنحنى Y كانت ٠,٧٥، و٠,٢٥، على التوالي. ومن الملاحظ أن b هي قيمة θ (النقطة على المحور الأفقي) الموافقة لـ ٠,٥ = p(θ)، وأن a هي ميل المنحنى في النقطة لـ ٥ = p(θ)، وأما في النموذج ثلاثي المؤشرات فإن b هي قيمة p(θ) الموافقة لـ ٥ (c + 1)، حيث c أو مؤشرالتخمين الزائف هي النقطة التي يتقاطع عندها منحنى الإجابات مع المحور الرأسي. ولعله من المفيد رسم عدد من منحنيات الإجابات الإحادية والثنائية والثلاثية المؤشرات باستخدام قيم المؤشرات التي يوفرها برنامج C-4 الموجود على القرص المرافق لهذا الكتاب.

الشكل ٢-٣

منحنيان لإجابات فقرات اختبار ذات مؤشرات مختلفة. (راجع النص لمزيد من المعلومات)



وقد تم التعبير عن مجال القدرة الكامنة في هذه الأشكال التوضيحية عن طريق وحدات الدرجات المعيارية (z)، إلا أن معظم التطبيقات التعليمية تحول الدرجات (z) إلى مقياس وسطه الحسابي ٣٠٠ وانحرافه المعياري ٥٠. وتساعد الرسوم البيانية (كالشكل ٢-٣) على رسم الخط البياني لنسبة الإجابات الصحيحة الموافقة لمستويات قدرة معينة كما يمكننا أن نرسم كذلك نسبة عدد الأفراد المنتمين لمستوى قدرة معين والذين إختاروا مموه معين مما يساعد في تقييم فعالية المموهات المرتبطة بفقرة اختبارية ما.

وفي الواقع العملي، تبقى مؤشرات الفقرة الاختبارية ودرجات مجال القدرة الكامنة (θs) للمتقدم للاختبار، مجهولة وتبقى المشكلة هي تحديد منحنى الإجابة الذي يتناسب بالشكل الأمثل مع الإجابات على الفقرات الاختبارية المختلفة. ويتطلب حل هذه المشكلة إجراءاً تكرارياً للاحتمالية القصوى يفترض عدداً من القيم الابتدائية لمؤشرات الفقرات الاختبارية ثم يحسب P(θ)s لمختلف قيم θ مع مقارنة الإجابات المتوقعة مع الإجابات الفعلية للفقرة، ويستمر إتباع الإجراء ذاته حتى يتم التوصل إلى الحل الأنسب. وتتطلب عملية تقديرمؤشرات الفقرات الاختبارية عدداً كبيراً من الإجابات لعدد كبير من الممتحنين يمثلون الشريحة السكانية التي ينتمي إليها كل من يُحتمل اختبارهم في المستقبل، ويتراوح هذا العدد بين ٥٠٠ للنموذج أحادي المؤشر (أو نموذج راش) و٢٠٠٠ للنموذج ثلاثي المؤشر.

ومن أهم خصائص مؤشرات الفقرات الاختبارية التي تم حسابها كون هذه المؤشرات مستقلة نوعاً ما عن مستوى القدرة المرتبط بالعينة التي تم على أساسها تحديد المؤشرات، إذ، وعلى عكس طرق الاختبار التقليدية التي تحصر درجة صعوبة الاختبار وتمييزه ضمن حدود العينة التي تم اختبارها، تعد هذه المؤشرات في نظرية الإجابة على فقرات الاختبار على المستوى النظري على الأقل مستقلة (أو لا علاقة لها) بالعينة موضع الاختبار وهذا ما يسمى بثبات (أو عدم تباين) العينة. ويساعد هذا الثبات لمؤشرات فقرات الاختبار على تحديد معدلات الدرجات المكافئة أو المعادلة في اختبارات مختلفة. وتتألف عملية المعادلة بين الاختبارات لتحويل درجات اختبار ما (أو صورة مختلفة من الاختبارذاته) إلى عدد من الوحدات مساوٍ لعدد وحدات اختبار آخر من الخطوات التالية: (١) إيجاد الرتب المئينية الموافقة لدرجات كلا الاختبارين، ومن ثم (٢) وضع جدول لدرجات مكافئة تتوافق مع الرتب المئينة في كلا الاختبارين.

وتتطلب عملية المعادلة في إطار (IRT) تحديد معادلة خطية تحول مؤشرات الصعوبة والتمييز لنموذج اختباري ما إلى نموذج اختباري آخر. وتسمى عملية إيجاد قيم ثابتة لمعادلة التحويل الخطية تضمن كون المؤشرات في كلا الاختبارين على نفس المقياس بعملية الربط. وتتطلب عملية الربط هذه أن يشترك الاختباران ببعض الفقرات الاختبارية (الفقرات الراسية Anchor items) أو أن يتم اختبار مجموعة فرعية من الأفراد بالاختبارين موضع الدراسة أو باختبار ثالث يقيس الخاصية ذاتها. كما ينطوي إجراء IRT لمعادلة الاختبارات، وهو إجراء اقتصادي، على اختبار مجموعات عشوائية من الأشخاص باستخدام مجموعات فرعية من الفقرات (اختيار الفقرات).

وتسمح عملية تحليل فقرات الاختبار (IRT) بالإضافة إلى تقدير مؤشرات الفقرات الاختبارية بتقدير درجات الممتحنين بالنسبة إلى مجال القدرة الكامنة. وفي الواقع يعتبر تحديد موقع الممتحن بالنسبة إلى مجال القدرة الكامنة الهدف الرئيس لنظرية (IRT). وكما كان الحال في تقدير مؤشرات الفقرة الاختبارية بناءاً على مستويات القدرة، فإن تحديد معدلات درجات القدرقهي عملية تكرارية تبدأ بتعويض قيم تجريبية لقدرة الممتحن ومؤشرات الفقرة التقديرية في المعادلة اللوجستية المناسبة. ثم تقارن قيم P(θ)s التقديرية مع قيم P(θ)s الفعلية وتستمر العملية حتى الوصول إلى المعادلة الأكثر تناسباً مع القيم الناتجة. كما يمكن حساب الأخطاء المعيارية للقيم المقدرة θs (وهي مقياس لتغير قيم θs التقديرية بالنسبة إلى قيم θs الفعلية المجهولة).

وتختلف IRT عن نظريات القياس التقليدية (التي تطبق الخطأ المعياري للقياس على جميع درجات مكونات عينة ما) بكون الخطأ المعياري للقياس يختلف من معدل درجات إلى آخر. ويساوي الخطأ المعياري لقياس مستوى قدرة معينة في نظرية IRT مقلوب كمية المعلومات التي تعبر عنهاالدرجات ضمن ذلك

المستوى. وتحدد كمية المعلومات الناتجة عن فقرة اختبارية ما عن طريق دلالة معلومات هذه الفقرة بينما يعبر مجموع قيم معلومات الفقرات الاختبارية المختلفة ضمن مستوى قدرة معين عن معلومات الاختبار بكامله ضمن هذا المستوى.

ويعد ثبات قدرة الممتحنين بالنسبة إلى الفقرات الاختبارية التي تم استخدامها لتحديد هذه القدرة واحداً من الصفات الهامة لنظرية IRT والتي تستند إلى عملية تقدير θs وتعني هذه الميزة أنه يمكن استخدام أي اختبار على أي درجة من الصعوبة لتحديد موقع الممتحن ضمن مجال القدرة الكامنة. إلا أنه وللحصول على أكثر النتائج دقة، ينصح باستخدام فقرات اختبارية (أو اختبار كامل) تتوافق درجة صعوبتها مع مستوى قدرة الممتحن.

هذا ويمكن استخدام IRT لتحقيق الأهداف التالية:

- وضع وتصميم الاختبارات.
- معايرة درجات الاختبارات للتوصل إلى مرجعية يمكن تفسير الدرجات هذه على أساسها.
- قياس الاختبارات.
- تحديد فعل الفقرات التبايني (DIF).
- الاختبار التكيفي.

تساعد نظرية IRT فيما يخص وضع وتصميم الاختبارات في تحديد مواقع تحليل الفقرات الاختبارية التي تعطينا قياسات أكثر دقة مما يقودنا إلى التركيز على هذه المواقع وعدم هدر الفقرات الاختبارية على مواقع أقل أهمية. وعليه يمكننا وباستخدام IRT تصميم اختبارات انتقاء شديدة الدقة واختبارات انتقائية، واختبارات ذات مرجعية محكِّية واختبارات تقليدية ذات طيف واسع الامتداد

ضمن مجال القدرة. أما مقاربة IRT في مجال تحديد عمل الفقرات التبايني (DIF)، فتتم برسم المخططات البيانية لمنحنيات الإجابة لكل مجموعة سكانية على حدة (البيض والسود، الرجال والنساء... إلخ). وتعبر الأشكال المختلفة لمنحنيات الإجابة لكل من المجموعتين عن عمل فقرات الاختبار التبايني (DIF).

اثبتت الاختبارات التكيفية (أو الاختبارات التي تتناسب مع خصوصية شخص ما) أنها نوع اقتصادي ودقيق من الاختبارات المحوسبة. وتعتمد هذه الاختبارات على قدرة الممتحن التي يتم تقديرها بناءً على إجابته على الأسئلة السابقة، وبما أن أدق الاختبارات هي تلك التي تتناسب صعوبة فقراتها مع مستوى قدرة الممتحن فإن عملية إعادة تقدير قدرة الممتحن التي تتكرر خلال الاختبار تسمح باختيار فقرات الاختبار الأكثر تناسباً مع مستوى قدرة الممتحَن الفعلية. وعلى الرغم من كون الاختبارات التكيفية أقصر من الاختبارات التقليدية إلا أن الاختبارات التي تتناسب مستوياتها مع مستوى قدرة الممتحَنين يمكن أن تكون أكثر صدقاً من الاختبارات الأكثر طولاً والمصممة لاختبار القدرة ذاتها. وقد تم اعتماد الاختبار التكيفي المحوسب (CAT) في عدد من الاختبارات القياسية مثل بطارية اختبارات القابلية المهنية للقوات المسلحة (ASVAB) واختبار القدرة الدراسية (SAT) واختبارات سجل الخريجين (GRE).

وتعد الحسابات المستخدمة في تحليل فقرات الاختبار (IRT) حسابات معقدة لا يمكن أن تتم دون الإستعانة بعدد من برامج الكمبيوتر مثلLOGIST وBILOG (ميسليفي وستوكينغ، ١٩٨٩) وASCAL (فايل، ١٩٨٥) وBIGSTEPS (رايت ولينأيكر، ١٩٩١). ولعل واحدة من المشكلات التي تعاني منها جميع نماذج الـIRT هي الافتراض بأن قدرةً كامنةً واحدة فقط هي المسؤولة عن الأداء الاختباري. وقد أحرزت بعض النماذج المتعددة الأبعاد نوعاً من التقدم في هذا السياق مثل نماذج (أمبرتسون، ١٩٩١ وريكيس، آكرمان وكارلسون، ١٩٨٨).

وتعتمد معظم نماذج الـ IRT على نظام توزيع درجات الفقرات (١-٠) وقد جرى العمل على نماذج أكثر تعقيداً تعتمد على نظام توزيع درجات متعددة (كقوائم الرصد) (مثال: ثيسين وشتاينبيرغ، ١٩٨٦؛ رايت وماسترز، ١٩٨٢).

الخلاصة

يتم القياس على عدة مستويات هي: المستوى الإسمي والمستوى الرُتبي والفئوي والنسبي. وتقع معظم طرق القياس النفسية ضمن المستوى الفئوي أو ما دونه. يتم تحليل مقاييس الأحداث الإدراكية والعاطفية والنفسية - الحركية بواسطة طرق إحصائية. وتتوفر إجراءات إحصائية أحادية المتغير (univariate) أو متعددة المتغيرات (multivariate) تصف وتحلل العلاقات والاختلافات بين المتغيرات النفسية. وتستخدم طرق الإرتباط كمعامل الإرتباط المولد للعزوم والتحليل العاملي والتحليل التمييزي بشكل كبير في عملية تفسير درجات الاختبارات.

وتركز نظرية الاختبار التقليدية على الصدق والثبات وقياس الاختبارات وتحليل فقرات الاختبار. ويعرّف الثبات على أنه قدرة الاختبار على قياس الشيء ذاته بالشكل ذاته في كل مرة يجرى فيها هذا الاختبار، بينما يشير الصدق إلى قدرة الاختبار على قياس الشيء الذي أُعلن أن هذا الاختبار يريد أن يقيسه، ويحدد صدق اختبار ما بواسطة الطرق التالية: صدق المحتوى والصدق المرتبط بالمحك (الصدق التلازمي والصدق التنبؤي) وصدق البناء. وتعد الإجراءات الإحصائية التي تقيِّم الصدق التنبؤي أو قدرة الاختبار على التنبؤ بالسلوك البشري إجراءات على درجة عالية من التطور.

ترى نظرية الثبات التقليدية أن المعدل الملاحظ لدرجات فرد ما في اختبار ما على أنه المجموع الفعلي لدرجات هذا الفرد مضافاً إليه خطأ معين في

القياس، وقد استندت الطرق التقليدية لتقييم الثبات (وهي طرق: إعادة الاختبار والصور المكافئة والإتساق الداخلي) إلى هذه المعادلة. تتطلب عملية حساب ثبات الإتساق الداخلي (وهي عملية بسيطة التكاليف) حساب معاملات التجزئة النصفية وكودر ريتشاردسون وكرونباخ. ويرتبط مفهوم الخطأ المعياري للقياس بالموثوقية إذ يُستخدم هذا المفهوم لحساب مجال القيمة العددية التي يمكن لنا القول ونحن على درجة محددة من الثقة إن معدل درجات الفرد الحقيقي يقع ضمنها. ويختلف ثبات الاختبار شأنه كشأن الصدق تبعاً للمجموعة قيد الاختبار والظروف المحيطة، وتؤكد نظرية التعميم على وجود عدة أنواع للثبات المرتبط بالاختبارات.

وقد قاد التركيز التاريخي على عمل الاختبارات في قياس الفروقات بين الأفراد إلى الاهتمام بقياس الاختبارات وتحديد المعايير التي يمكن تفسير نتائج الاختبارات على أساسها. ومن أهم هذه المعايير: معيارالعمر والصف الدراسي والمعايير المئينية ومعايير الدرجات المعيارية. وتعد المعايير الثلاثة الأولى مقاييساً على المستوى الرُتبي وهي المقاييس المفضلة للمدرسين وغيرهم من الذين لا يتقنون فنون القياس النفسي الأكثر تعقيداً. أما محترفو القياس النفسي فيفضلون معيار الدرجات المعيارية الذي يقيس على المستوى الفئوي. وتعتمد درجة الفائدة التي يمكن الحصول عليها من مجموعة من معايير الاختبارات على ما إذا كانت هذه المعايير قد تم الحصول عليها من عينة (ويفضل أن تكون عينة عشوائية متعددة الطبقات) تمثل المجموعات البشرية التي يستهدفها الاختبار.

تعتبر عملية تحليل عمل الفقرات الاختبارية خطوةً هامة في مجال تطوير الاختبارات، ويعتمد هذا التحليل على مقارنة الإجابات على الفقرات الاختبارية المختلفة إما مع مَحكَّات داخلية (كدرجات الاختبار الكلية) أو محكَّات خارجية (كالتحصيل الأكاديمي والأداء الوظيفي). ويحسب في حالة اختبارات الخيارات

المتعددة، بالإضافة إلى مؤشري الصعوبة والتمييز، توزيع إجابات المموهات وذلك لتحديد ما إذا كانت هذه المموهات تعمل بالشكل المطلوب. ويتم، بناءً على عملية تحليل فقرات الاختبارات حذف أو مراجعة الفقرات غير الفعّالة، كما يتم في حالة الاختبارات القياسية تحديد ما إذا كانت هذه الاختبارات تعمل بشكل مختلف عند اختلاف المجموعة السكانية موضع الاختبار. وقد تم بحث إجراءات عمل الفقرات الاختبارية التبايني (DIF) في هذا الفصل من الكتاب.

وتهدف نظرية IRT على عكس نظريات الاختبار التقليدية التي تركز على مجموع درجات الاختبار إلى تحليل الإجابات على الفقرات الاختبارية المحددة وربطها بموقع الممتحن بالنسبة إلى مجال القدرة الكامنة. وبناءً على النموذج الرياضي (المنطقي أو العادي) الذي يربط درجات القدرة مع احتمال الإجابة على فقرة اختبارية ما بشكل صحيح، يمكّننا الإجراء الإحصائي التكراري IRT من تقدير درجات القدرة الكامنة بالإضافة إلى المؤشرات المناسبة (كالصعوبة والتمييز والتخمين الزائف) للنموذج الاختباري. ومن أهم مميزات IRT كون نماذجه مستقلة عن الاختبار وعن الفقرة الاختبارية (على المستوى النظري على الأقل).

يمكن استخدام IRT لتحقيق أغراض أخرى مثل وضع وتصميم الاختبارات ومعادلة الاختبارات ومعايرة الفقرات الاختبارية وتقييم عمل الفقرات الاختبارية التبايني (DIF) والاختبار التكيفي.

مصطلحات للمراجعة:

راجع معاني المصطلحات المدرجة أدناه والتي تم استخدامها في هذا الفصل. الرجاء مراجعة فهرس التعريفات أو المعجم في حال وجود أي التباس حول معاني المصطلحات.

الاختبار التكيفي	التحليل العاملي
معيار العمر	الخطأ السلبي غير الحقيقي
فقرات الاختبار	الخطأ الإيجابي غير الحقيقي
النسبة القاعدية	معامل التعميم
الإرتباط القانوني	نظرية التعميم
المعامل ألفا	معيار الصف الدراسي
معامل التكافؤ	صدق الإتساق الداخلي
معامل الثبات	القياس على المستوى الفئوي
معامل الثبات والتكافؤ	تحليل فقرات الاختبار
الاختبار التكيفي المحوسب	معامل السهولة والصعوبة
الصدق التلازمي	معامل التمييز
صدق البناء	نظرية استجابة الفقرة (IRT)
صدق المحتوى	معادلات كودر- ريتشاردسون
المصادقة عن طريق التقارب	الارتباط
الاختبار ذو المرجعية المحكية	القياس

صدق المحك	معيار العمر العقلي
عمل الفقرة الاختبارية التبايني (DIF)	الانحدار المتعدد
التحليل التمييزي	المستوى الاسمي للقياس
المصادقة على المميز	الدرجات المعيارية المطبّعة
المظاهر	الاختبار ذو المرجعية المعيارية
التعريف العملياتي	معدل الاختيار
العملياتية	طريقة التجزئة النصفية (ثبات نصف الاختبار)
المستوى الرُتبي للقياس	الخطأ المعياري للتقدير (التخمين)
معامل الصورة الاختبارية المكافئة	الخطأ المعياري للقياس
تحليل المسار	الدرجات المعيارية
تقنين الاختبارات	المعيار المئيني
العينة العشوائية متعددة الطبقات	الصدق التنبؤي
جدول المواصفات	القياس النفسي
معامل إعادة الاختبار	نموذج راش
	مؤشر النسب
ثبات العينة	الثبات

المراجع المقترحة:

Drasgow, F., & Hulin, C. L. (1990). Item response theory. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed.) (Vol. 1, pp. 577–636). Palo Alto, CA: Consulting Psychologists Press.

Embretson, S. (1993). Psychometric models for learning and cognitive processes. In N. Frederiksen, R. J. Mislevy, & I. I. Behar (Eds.), *Test theory for a new generation of tests* (pp. 125–150). Hillsdale, NJ: Erlbaum.

McKinley, R. L. (1989). An introduction to item response theory. *Measurement and Evaluation in Counseling and Development, 22*, 37–57.

Shavelson, R. J., Webb, N. M., & Rowley, G. L. (1989). Generalizability theory. *American Psychologist, 44*, 922–932.

Thorndike, R. L., & Thorndike, R. M. (1994). Reliability in educational and psychological measurement. In T. Husén & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (2nd ed.) (Vol. 9, pp. 4981–4995). Tarrytown, NY: Elsevier.

Weiss, D. J., & Yoes, M. E. (1991). Item response theory. In R. K. Hambleton & J. N. Zaal (Eds.), *Advances in educational and psychological testing: Theory and applications* (pp. 69–95). Boston: Kluwer.

Yen, W. M. (1992). Item response theory. In M. Alkin (Ed.), *Encyclopedia of educational research* (6th ed.) (Vol. 2, pp. 657–666). New York: Macmillan.

Zeller, R. A. (1994). Validity. In T. Husén & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (2nd ed.) (Vol. 11, pp. 6569–6576). Tarrytown, NY: Elsevier.



تصميم ووضع وتصحيح الاختبارات

تتوازى المعلومات عن تصميم الاختبارات الموجودة في القسم الأول من هذا الفصل مع معلومات الفصل الثاني من كتاب سلالم التقدير وقوائم الرصد (أيكين، ١٩٩٦) الذي يعالج قضية تصميم ووضع سلالم التقدير وقوائم الرصد. وكذلك مع معلومات الفصل الثاني من كتاب الاستبانات ومقاييس السمات (أيكين، ١٩٩٧). ومن باب التذكير، إن الكتابين المذكورين آنفاً هما الجزءان الأول والثاني من هذه السلسة. يناقش هذا الفصل أيضاً ثلاثة عناصر أخرى من عناصر الاختبار هي: إجراء الاختبار وتصحيحه ثم تفسير نتائجه. وتساعد برامج الفئة ب الموجودة على القرص المرافق لهذا الكتاب على فهم وإنجاز هذه المهام، إذ يساعد برنامج B-1 على تصميم اختبارات موضوعية، ويساعد برنامج B-2 على إجراء ووضع درجات الاختبارات الموضوعية، بينما يساعد برنامج B-3 على عملية "التصحيح في حالة التخمين" Guessig في الاختبارات الموضوعية. أما برنامج B-4 فيساعد على وضع درجات أسئلة المطابقة Matching أو إعادة الترتيب. ويستخدم برنامج B-5 في وضع درجات الأسئلة التي تتطلب ترتيب

الفقرات. أما برنامج B-6 فيستخدم في وضع الدرجات وفق طريقة كاجوري المعدلة (Cajori)، بينما يفيد برنامج B-7 والذي يعطي تقديراً لاحتمالية إرتكاب ممتحنين جالسين على مقربة من بعض نفس الأخطاء في التعرف على حالات الغش في الاختبارات.

تصميم الاختبارات

لا تتطلب الغالبية الساحقة من الاختبارات النفسية والتعليمية تصميماً شديد الدقة أو إعداداً شاقاً وصعباً إذ أن معظم هذه الاختبارات هي اختبارات وليدة اللحظة يضعها المدرسون وغيرهم من المقيمين الذين هم بحاجة إلى وسيلة مختصرة وبسيطة تمكنهم من تقييم تحصيل الفرد العلمي أو مكانته. وغالباً ما تستأثر عملية التدريس بجل اهتمام المدرسين فلا يعيرون اهتماماً مناسباً لعملية قياس ما إذا كانت الأهداف التعليمية قد تم تحقيقها أوإلى أي درجة تم تحقيقها. وعليه، لا يقضي هؤلاء المدرسون عادةً وقتاً طويلاً في وضع اختبارات مقالية أو اختبارات موضوعية تقيس درجة تحصيل الطلاب. وعلى عكس الاختبارات التي يضعها المدرسون والتي تعد اختبارات "الفرصة الواحدة" للطالب، تعد الاختبارات وأدوات التقييم التي يضعها محترفو تصميم الاختبارات والتي تسوقها شركات تجارية للاختيارات.

يتوفر عدد كبير من الاختبارات التي تقيم المهارات الإدراكية والعاطفية والنفسية – الحركية، كما يتم تصميم عدد كبير من الاختبارات كل يوم والتي قد تنشر أو لا تنشر. وتختلف طرق تصميم الاختبارات إلى حد ما في نوع الاختبار ولكن في الإحالة.

يجب أن يبدأ الاختبار ببيان الفرض من الاختبار، ويتضمن هذا البيان بشكل عام إجابات للأسئلة التالية:

١- ما هي المتغيرات التي يهدف الاختبار قيد التصميم إلى قياسها؟

٢- من سيتم اختباره؟

٣- ما هي الظروف التي سيجرى ضمنها الاختبار؟

٤- كيف سيتم وضع درجات الاختبار؟

٥- ما هي الجهة التي سوف يتم تبليغها بنتائج الاختبار وكيف سيتم استخدام هذه النتائج؟

٦- ما هي الموارد المالية المتوفرة لدعم عمليات تصميم وإجراء الاختبارات ثم وضع درجاتها وتحليل نتائجها؟

وللإجابة على السؤال الأول، يتوجب على مصمم الاختبار أن يحدد ما إذا كان المطلوب هو قياس التحصيل أو الذكاء أو الاستعدادات الخاصة، إذ غالباً ما تركز الاختبارات التحصيلية على ما قد تعلمه الممتحن ضمن مجال مهاري أو معرفي معين في حين تتطلع اختبارات الاستعداد إلى المستقبل أو ما يمكن لفردٍ ما أن يتعلمه إذا توفرت له الظروف الموائمة. ولهذا السبب تتناول معظم فقرات الاختبارات التحصيلية مادةً علمية محددة تم تدريسها بشكل تقليدي ولا تتطرق إلى التجارب الإنسانية العامة. ويجب عند تصميم الاختبارات التحصيلية التي تختبر مواد علمية محددة كالقراءة والرياضيات واللغة الانجليزية أو اللغات الأجنبية إعداد قائمة بمواصفات المادة العلمية موضع الاختبار، ثم يجب إتخاذ قرارات تتعلق بأنواع الفقرات الاختبارية التي يمكنها أن تمثل المادة العلمية بشكل جيد وتقيسها بشكل فعال.

وتفيد الإجابة على السؤالين ٢ و٣ المتعلقين بمن سوف يتم اختباره وما هي الظروف التي سيجري ضمنها الاختبار في تحديد نوعية الفقرات الاختبارية المستخدمة (فقرات موضوعية أو مقالية أو أدائية) في الاختبار وتحديد طول

الاختبار وشكل تقديم الاختبار (على شكل كتيب أو تجربة عملية أو مقاطع شفهية) واللغة المستخدمة في الفقرات المطبوعة أو المحكية. وغالباً ما تكون الاختبارات الفردية مؤلفة من عدد من وسائل الاتصال (شفهية وكتابية وتطبيقية) التي تتطلب تعليمات شفهية مختلفة لكل قسم أو مجموعة من الفقرات. أما الاختبارات الجماعية فيتم تقديمها على شكل كتيب وترافق التعليمات المكتوبة في الكتيب تعليمات شفهية بسيطة من القائمين على الاختبار. ترتبط القرارات المتخذة بشأن محتوى الاختبار وشكله على الوقت المتوفر لإجراء الاختبار، إذ يمكن خلال ثلاث ساعات طرح عدد أكبر من الأسئلة واختبار عدد أكبر من المتغيرات مما يمكن طرحه خلال ساعتين إلا أن المشكلة تكمن في كون الاختبارات التي تستمر لأكثر من ساعتين اختبارات تؤثر سلبياً على تركيز الممتحن ورغبته في الإجابة على الأسئلة الأمر الذي يحول دون قياس قدرة الممتحن الحقيقة بشكل دقيق. ويتناول هذا الفصل من الكتاب الطرق المحوسبة في إجراء الاختبارات التي يجب أن تؤخذ بعين الاعتبار عند تصميم الاختبارات.

وتلعب طريقة وضع درجات الاختبار (باليد أو عن طريق الماسحة الالكترونية البصرية أو عن طريق إعطاء الممتحن فرصة وضع درجاته بنفسه) دوراً هاماً في عملية تصميم الاختبارات. ففي معظم الاختبارات الفردية يكون القائم على الاختبار وواضع الدرجات (أو المصحح) هما الشخص ذاته وغالباً ما تتم عملية التصحيح أثناء عملية الاختبار، أما في حالة الاختبارات الجماعية فلا يتم وضع الدرجات إلا بعض الانتهاء من الاختبار بكامله حيث يقوم القائم على الاختبار أو موظف مسؤول أو أي شخص كفؤ بوضع الدرجات. وتعتمد درجة تفصيل المعلومات المكتوبة في كشف الدرجات على الجهة المعنية بقراءة هذه المعلومات وعلى طريقة الحصول عليها. إذ يجب أن تتم كتابة المعلومات بطريقة مفهومة

لكل من لديه الحق بالإطلاع على النتائج ولديه القدرة على استخدام المعلومات الناتجة لما فيه مصلحة الممتحن.

ويمكن طرح عدد من الأسئلة المرتبطة بالموارد المالية المتوفرة لتصميم الاختبار: هل سيتقاضى مصمم الاختبار أجراً وما هو مقدار هذا الأجر؟ هل من الأفضل والأوفر تقديم الاختبار في قالب معين أو قالب آخر؟ ما هي الظروف المثلى اقتصادياً لإجراء الاختبار وتصحيحه؟

التخطيط وتحديد مواصفات الاختبارات التحصيلية

تختلف خطة ومواصفات الاختبارات التحصيلية عن خطط ومواصفات اختبارات القدرات الإدراكية العامة أوالخاصة. وتختلف مواصفات اختبار التحصيل باختلاف نوع الاختبار من حيث كونه اختبار خلاصي (أو اختبار نهائي يأتي في نهاية الوحدة الدراسية أو المقرر بأكمله) أو اختبار تكويني Formative (أي اختبار يجرى دورياً يتم على أساسه تعديل المقرر أو طريقة التدريس). وعلى الرغم من أن مواصفات الاختبار تختلف باختلاف موضوع الاختبار، يبقى إعداد جدول بمواصفات مفصلة لمحتوى الاختبار أمراً لا بد منه لضمان تحقيق الاختبار لأهدافه.

درج الباحثون في مجال التعليم خلال السبعينيات والستينيات على وضع أهداف محددة بدقة جداً للمقررات التعليمية وقد تم استخدام دلائل التصنيف مثل التي وضعها بلوم وكارثول (١٩٥٦) أو غرلاش وسوليفان (١٩٦٧) أو إيبيل (١٩٧٩) في تصميم أدوات التعليم والتقييم. ويتألف دليل بلوم وكارثول (١٩٥٦، ١٩٨٤)، وهو الدليل الأكثر شعبية للأهداف التعليمية، من ترتيب هرمي ينقسم إلى ٦ مجموعات هي: المعرفة والفهم والتطبيق والتحليل والتركيب والتقييم. لا تعد هذه المجموعات الستة التي تم ترتيبها من الأبسط إلى الأعقد ومجموعات

غير منتمية ولكنها شاملة بشكل متبادل بحيث أن الأهداف التعليمية المنتمية إلى مجموعات في الطبقات الدنيا تم تضمينها في مجموعات الطبقات العليا بالضرورة.

تم استخدام دلائل التصنيف مثل تلك التي وضعها بلوم وكارثول (١٩٥٦) لتحديد أحد أبعاد جداول مواصفات الاختبارات. وتُصنف أهداف المقرارات أو الوحدات التدريسية غالباً على أنها جداول ذات بعدين، يتألف البعد الأول (الحقول الأفقية) منها من مجموعات الأهداف السلوكية والبعد الثاني من مجموعات أهداف محتوى الاختبار (الحقول العمودية). فعلى سبيل المثال، يمكن لجدول مواصفات اختباريقييم عمليتي إعداد وإجراء الاختبارات التحصيلية أن يتألف من أربعة حقول أفقية تحدد الأهداف السلوكية للاختبار (معرفة المصطلحات، معرفة الحقائق المحددة، الاستيعاب، التطبيق) وخمسة أعمدة لمواصفات المحتوى (الإعداد للاختبار، وضع الاختبار، إجراء الاختبار، وضع الدرجات، تحليل فقرات الاختبار). وعند نقطة تقاطع الحقول الأفقية والعمودية، تتوفر قوائم تحتوي وصفاً مختصراً لفقرات الاختبار التي تتناول موضوعاً ما وذكراً لعدد هذه الفقرات.

وعلى الرغم من أن معظم المدرسين وغيرهم من واضعي الاختبارات من غير المحترفين لا يضعون جدولاً مفصلاً بمواصفات الاختبارات الخاصة بظرف ما التي يستخدمونها، يبقى من الحكمة أن يكون لدى واضع الاختبار خطة دقيقة للاختبار وغالباً ما تحتوي الخطة التفصيلية لمحتوى الاختبارعلى قائمة بالمجموعات الأساسية والفرعية لكل موضوع يجب أن تتم تغطيته وعدد (أو نسبة) الفقرات الاختبارية التي تم تخصيصها لكل من هذه المواضيع. وبهذه الطريقة يضمن واضع الاختبار تغطية الاختبار لكافة المواضيع المراد اختبار تحصيل الطلاب فيها كما يضمن واضع الاختبار أن الاختبار مناسب من حيث الطول والشكل للممتحنين وللوقت والموارد المتاحة. ومن الحكمة إعداد فقرات

اختبار تزيد ٢٠٪ عن العدد المطلوب للاختبار في شكله النهائي، ويتم الاختيار النهائي للفقرات بعد اتمام عمليات المراجعة وبعد الإطلاع على نتائج تجربة الفقرات الاختبارية على عينة من الأشخاص تمثل المجموعة البشرية المراد اختبارها لاحقاً.

تحضير فقرات الاختبار

يبحث الفصل الخامس من هذا الكتاب إجراءات وضع فقرات الاختبار الشفهية واختبارات الأداء بينما تتناول الفصول ٦ -٨ فقرات الاختبار المناسبة للاختبارات بشكل عام ولاختبارات الاستعداد بشكل خاص. وتنطبق العديد من التوصيات المتعلقة بوضع فقرات اختبارات التحصيل على أنواع أخرى من المقاييس الإدراكية.

ويمكن تقسيم فقرات الاختبار المستخدمة في تلك التي تستخدم فيها الورقة والقلم إلى الفئات التالية: الاستذكار بالمقابل مع الإدراك، أو الاستجابة المركبة بالمقابل مع التعرف على الأشياء، أو الفقرات المقالية مقابل الأسئلة الموضوعية. وتتطلب الأسئلة المقالية من الممتحن استذكار المعلومات المطلوبة من الذاكرة أو تزويد القائم على الاختبار بهذه المعلومات أو تركيب هذه المعلومات، في حين تتطلب الإجابة على الأسئلة الموضوعية عمليات إدراك الإجابات الصحيحة أو التعرف عليها أو اختيارها. وقد ازدادت شعبية فقرات الاختبارات الموضوعية بسبب سهولة تصحيحها وكونها أكثر موضوعية من الأسئلة المقالية. ومن الضروري لفت نظر القارئ هنا إلى أن فقرات الاختبار الموضوعية والمكتوبة بشكل محكم تعطي الممتحن فكرة أكثر وضوحاً من الأسئلة المقالية عن المهمة التي يجب إنجازها لاجتياز الاختبار بنجاح، إلا أن الموضوعية الفعلية لهذا النوع من الأسئلة لا تكمن في الأسئلة بحد ذاتها إنما في عملية وضع الدرجات. وهذا

يعني أن مصححَين مختلفين سوف يضعان الدرجات نفسها للاختبار نفسه، على حين يصعب حدوث هذا في حالة الأسئلة المقالية.

فقرات الاختبار المقالية

على الرغم من كون الاختبارات الموضوعية مقاييس فعالة للذاكرة الإدراكية إلا أن هذه الاختبارات تنقصها القدرة على قياس العمليات الإدراكية المعقدة كالقدرة على التنظيم والسرد والتواصل. تقيس الأسئلة المقالية هذه العمليات كما أنها تتطلب وقتاً أقل للإعداد ولا تترك للمجيبين عليها مجالاً للتخمين. إلا أن هذه الأسئلة، ومع الأسف، غالباً ما تكون أسئلة عامة جداً مما يجعل معناها يختلف من شخصٍ إلى آخر. وعلاوةً على ذلك لا يمكن للممتحن الإجابة إلا على عدد محدود جداً من الأسئلة المقالية ضمن فترة الاختبار، مما لا يسمح بتغطية المعرفة العلمية التي سيتم تقييم الممتحنين على أساسها. أما الممتحنون الذين لا يملكون فكرة واضحة عن المادة العلمية موضع الاختبار فغالباً ما يكون تقييمهم مخادعاً. وأخيراً تبقى المشكلة الكبرى في الاختبارات المقالية كون هذه الاختبارات تتطلب وقت تصحيح طويل جداً مقارنة بالأسئلة الموضوعية.

وتتوفر عدد من التوصيات التي يمكن أن تجعل الاختبارات المقالية أكثر موضوعية وأكثر شمولاً:

١- تحديد هدف السؤال والتأكد من أن معناه شديد الوضوح مما يسمح للخبراء بالتوصل إلى القرارات ذاتها فيما يتعلق بكون إجابةٍ ما إجابةً جيدة أو سيئة.

٢- استخدام عدد أكبر من الأسئلة المقالية التي تتطلب إجابات قصيرة بدلاً من عدد أقل من الأسئلة ذات الإجابات الطويلة.

٣- تحديد الطول المطلوب للإجابة (صفحة أو نصف صفحة... إلخ).

٤- الطلب من الممتحنين الإجابة على جميع فقرات الاختبار والطلب منهم وضع إجابة كل فقرة على ورقة منفصلة.

فقرات الإجابات القصيرة

تقع الفقرات الاختبارية التي تطلب من الممتحن إجابات قصيرة كملء الفراغات أو الإجابة على سؤال بعدد من الكلمات أو العبارات بين الأسئلة ذات الإجابات المقالية الطويلةوأسئلة التعرف (كأسئلة الخيارات المتعددة وأسئلة الصواب أو الخطأ وأسئلة المطابقة). وعلى الرغم من سهولة وضع أسئلة الإجابات القصيرة، تعد هذه الأسئلة محدودة من حيث تقييمها لمعرفة المصطلحات أو الحقائق كما يمكن أن تقود إلى أخذ الإجابات من النص مباشرةً ويمكن أن تحتمل إجابتين أو أكثر مما يجعل هذه النوع من الأسئلة أقل موضوعية مما هو متوقع. ويجب أن تصاغ أسئلة الإجابات القصيرة بشكل لا يحتمل الغموض أو الالتباس، وتعد الأسئلة المباشرة أفضل بكثير من العبارات الناقصة، ويفضل وضع الفراغ في نهاية العبارة وليس في منتصفها. ويجب تفادي وضع الفراغات كلها في قسم محدد من السؤال ومن الأفضل أن تكون جميع الفراغات من الطول ذاته. وفي حال كون الإجابات المطلوبة إجابات عددية يجب أن يتم تحديد الوحدة المطلوبة للإجابة.

يحتوي برنامج F-1 الموجود على القرص المرن المرافق لهذا الكتاب على نموذج اختبارملء فراغات للعمليات الحسابية البسيطة (الجمع والطرح والضرب والتقسيم ورفع الأعداد المكونة من آحاد أو آحاد وعشرات إلى أس معي). ومن ميزات هذا البرنامج كون العددين الذين يولدهما البرنامج للمسائل المختلفة عددين عشوائيين مما يعني حصول كل ممتحن على مجموعة مختلفة من المسائل.

ويعد نمط أسئلة ملء الفراغات ضمن المقاطع النصية (أسئلة كلوز cloze) نمطاً مناسباً جداً لتقييم مهارات القراءة. إذ تعتمد هذه الطريقة على حذف عدد معين من الكلمات من نصٍ ما ويطلب من الممتحن أن يجد كلمات مناسبة لوضعها مكان ما تم حذفه. وتعد درجة قدرة الممتحن على فهم النص وملء الفراغات بكلمات مناسبة مقياساً لقدرات الممتحن في مجال القراءة. ومن الأمثلة على أدوات القياس هذه كتاب اختبارات النصوص الناقصة (هودر وستوفتون) التي تحتوي عادةً على ٥٠ فراغاً. وتفيد هذه الاختبارات ليس في قياس قدرة الأفراد على القراءة فحسب بل تفيد المعلومات عن مدى تمكن مجموعة من الأفراد من ملء فراغات النص بنجاح في تحديد ما إذا كان النص بحد ذاته مقروءاً (يسهل استيعابه) أم لا.

أسئلة الخطأ والصواب

على الرغم من أن أسئلة الخطأ والصواب هي أسئلة سهلة الإعداد إلا أنها غالباً ما تُستخدم بشكلٍ خاطئ. وتُمكن سهولة وسرعة قراءة أسئلة الصواب والخطأ الممتحنين من الإجابة على ضعف عدد أسئلة الخيارات المتعددة في نفس المدة الزمنية. إلا أن هذا النوع من الأسئلة، شأنه كشأن الأسئلة ذات الإجابات القصيرة، يتيح للممتحن فرصة الحصول على الإجابات مباشرة من النص كما أنه غالباً ما تكون الأسئلة سطحية وتتعامل مع أمور بدهية. ويتوجب على مصممي الاختبارات من نوع اختبارات الخطأ والصواب، تفادي استعمال أدوات الدلالة المحدِّدة Specific determiner مثل: غالباً أو أحياناً أو عادةً التي ترتبط عادةً بالعبارات الصحيحة وكلمات مثل: كل ودائماً وأبداً وفقط التي ترتبط بالعبارات الخاطئة عند وضع هذا النوع من الأسئلة. يمكن أن يقود استخدام أدوات الدلالة المحددة التي يمكن أن تضفي هالة من الصدق على المقولات الكاذبة واستخدام المقولات التي تحتوي على منطقٍ سطحي أو اعتقادات

خاطئة شائعة أو فيها صدى الحقيقة بعض الأشخاص غير المتمكنين إلى اختيار بدائل الإجابة الخاطئة. ويجب تفادي كذلك الجمل التي تحتوي على تصريحات عامة جداً أو التي تستخدم النفي (والنفي المزدوج بشكلٍ خاص) أو الجمل الشديدة التعقيد والغموض أو الجمل الطويلة المعقدة أو التي تحتوي عدداً من الأفكار أو آراءاً شخصية (إلا في حال ذكر المصدر الذي تم أخذ الرأي منه). وأخيراً يجب أن تتساوى أسئلة الصواب والخطأ في طولها وعددها. وتحتوي برامج F-3 وF-4 على اختبارين مؤلفين من ٥٠ سؤال صواب و خطأ يتناول الأول منهما موضوع الشيخوخة والتقدم في العمر بينما يتناول الثاني موضوع الموت. ويمكن إجراء الاختبارين وتصحيحهما عن طريق البرنامجين F-3 وF-4.

أسئلة المطابقة

تعد فقرات الاختبار التي تعتمد على وضع عمودين يحتوي الأول منهما على عدد من الفرضيات Premises (أو المحفزّات) بينما يضم الآخر الإجابات المناسبة والتي يطلب من الممتحن أن يتعرف عليها نوعاً من فقرات الاختبار القليلة الاستخدام والتي يندر وجودها في الاختبارات القياسية. وتعد فقرات الاختبار التي تعتمد على إعادة الترتيب نوعاً من فقرات المطابقة التي تتطلب من الممتحن ترتيب احتمالات الإجابة ضمن مجموعات مرتبة. ويراعى عند تصميم فقرات المطابقة ترقيم الفرضيات تسلسلياً وباستخدام الأعداد وترقيم اختيارات الإجابة بحروف متسلسلة. ويعد عدد من الفرضيات يتراوح بين ٦ إلى ١٥ عدداً معقولاً يرافق كل فرضية منها اختياران أو ثلاثة. ومن الضروري وضع الفقرة الاختبارية بكاملها على صفحة واحدة وتحديد الأساس الذي يتم الاستناد إليه عند المطابقة بين الفرضيات والإجابات. تعد فقرات المطابقة، بالرغم من كونها مقاييس لمعرفة الحقائق المحددة (راجع شاها، ١٩٨٤، لرأي مغايير)، فقرات اختبارية سهلة التصميم وتغطي المادة العلمية المراد اختبارها بشكل جيد بالمقارنة

مع غيرها من أنواع الفقرات الاختبارية. إن تحديد خيارات الإجابة بنفس النوع من المعلومات (أسماء، تواريخ، أماكن، إلخ) يعدّ محدداً رئيساً في اختبارات مطابقة الفقرات.

يمكن استخدام برنامج B-4 لإجراء وتصحيح اختبارات المطابقة وإعادة الترتيب، ويحتوي الملف "match.1" على العمود الأول أو الفرضيات بينما يحتوي الملف "match.2" على العمود الثاني أو عمود الإجابات، أما مفتاح الإجابات فيوجد في الملف "match.3" تحتوي الملفات أعلاه على الفرضيات والإجابات ومفتاح الإجابات لسؤال مطابقة يتناول تاريخ التقييم النفسي ويمكن استبدال محتويات هذا السؤال بأي معطيات أخرى للحصول على أسئلة مطابقة أو إعادة ترتيب جديدة.

أسئلة الخيارات المتعددة

تعد نماذج الاختيارات المتعددة أكثر صور الاختبارات الموضوعية تعدداً للاستعمال، وهي الشكل الأكثر استخداماً في الاختبارات الموحدة القياسية. تقيس هذه الفقرات المهارات المعقدة والبسيطة ويتم تصحيح هذه الفقرات بشكل أسهل وأسرع بكثير من الأسئلة المقالية. وبالرغم من إحتمالية إعطاء الممتحن للإجابة الصحيحة عن طريق التخمين، تبقى هذه الاحتمالية أقل منها في حالة أسئلة الصواب والخطأ. ويمكن الحصول على معلومات تشخيصية مفيدة عن طريق تحليل الإجابة للاختيارات الخاطئة (المشتّتات، المموهات). لا يخلو هذا النوع من فقرات الاختبار من العلل، ولعل من أهمها كون عملية وضع هذا النوع من الأسئلة تتطلب وقتاً وجهداً كبيرين. كما أن هذا النوع من الأسئلة يركز على مهارات التعرف بدلاً من مهارات التنظيم ويتطلب وقتاً أطول للإجابة من أسئلة الصواب والخطأ، كما يمكن أن لا يغطي هذا النوع من الأسئلة المادة العلمية موضع الاختبار وأن يعاقب الأفراد الذين يفكرون بعمق (راجع هوفمان، ١٩٦٢).

يحتوي الصندوق ٣-١ على قائمة توصيات مستقاة من أفضل الممارسات في مجال وضع أسئلة الخيارات المتعددة. إن برامج الحاسب التي تساعد على إعداد فقرات الاختبار (مثال: TUTOR ،PLANIT ،COURSEWRITER)، ومواصفات المواد والمقررات الدراسية الخوارزمية التي تساعد على توليد أسئلة اختبارات، وطرق كتابة فقرات الاختبار التي تستند إلى المهارات اللغوية ومعرفة المفاهيم، تَعِدُ بإجراءات أكثر فاعلية ودقة لإعداد أسئلة الخيارات المتعددة (هيرمان، ١٩٩٤).

وبرغم وجود قواعد وخوارزميات Algorithms لكتابة أسئلة الخيارات المتعددة، تبقى عملية الكتابة هذه نوعاً من الفن والعلم في آنٍ واحد. ويمكن تعقيد مهمة الممتحن باستخدام العبارات التالية: "كل ما ذكر أعلاه" و"لا شيء مما ذكر أعلاه" و"شيئان مما ذكر أعلاه" و "كل ما ذكر أعلاه إلا شيء واحد" و"واحد مما يلي." كما يمكن أن تكون جميع الإجابات صحيحة إلا أن على الممتحن اختيار الإجابة الأفضل أو الأكثر صحة. ومن المقترحات الأخرى لجعل الأسئلة متعددة الخيارات أكثر صعوبة: الطلب من الممتحن أن يتعرف على الإجابات الصحيحة من ضمن عدد متغير من الإجابات أو إجراء تحسينات (تعديلات) على الإجابات المعطاة أو التعرف على النص الصحيح لمسألة رياضية ما.

يحتوي الصندوق ٣-٢ على مقترحات أخرى لجعل اختيار الإجابات الصحيحة في أسئلة الخيارات المتعددة عملية تتطلب تحليل وتفكير ولا تقتصر فقط على الحفظ عن طريق التكرار.

الصندوق ٣-١

توصيات لإعداد الأسئلة المتعددة الاختيارات

١- اجعل جذر السؤال إما جملة غير مكتملة أو سؤال (وهذا الأخير هو الخيار الأفضل). وفي حال استخدام جملة غير مكتملة يجب وضع الفراغ في نهاية الجملة.

٢- تأكد من صياغة جذر السؤال صياغة واضحة ومتناسبة مع إمكانيات الممتحنين اللغوية.

٣- لا تكثر من استخدام أسئلة الآراء وفي حال استخدام هذا النوع من الأسئلة أشر بوضوح إلى مصدر الرأي.

٤- حاول أن يشتمل جذر السؤال على أكبر قدر ممكن من الفقرة الاختبارية بما فيه العبارات المتكررة التي تنطبق على جميع البدائل (الاختيارات).

٥- اجعل الاختيارات كلها متساوية الطول وصحيحة من حيث ارتباطها النحوي مع جذر السؤال ومعقولة لمن لا يعرف الإجابة على السؤال.

٦- حدد، عند صياغة المموهات، الأسباب التي يمكن على أساسها أن يختار الممتحنون الذين يجهلون الإجابة الصحيحة هذه المموهات.

٧- الأخطاء الشائعة التي تعد صحيحةً جزئياً هي نوع جيد من المموهات.

٨- تأكد من عدم وجود أية إشارات لغوية أو أي أدلة ضمن جذر السؤال تشير إلى الإجابة الصحيحة.

٩- على الرغم من أن معظم أسئلة الخيارات المتعددة تضم أربعة أو خمسة خيارات، يمكن الإكتفاء بخيارين أو ثلاثة وخصوصاً في المراحل الدراسية الدنيا.

١٠- رتب الخيارات التي تتطلب ترتيباً بطبيعتها (الأعمار أو التواريخ) تبعاً لترتيبها الطبيعي أو بشكل عشوائي أو وفق الترتيب الهجائي.

١١- حاول تفادي (أو الإقلال من) استخدام أدوات النفي (ليس) أو أدوات الدلالة المحددة (دائماً وأبداً) و جمل مثل: "كل ما ورد أعلاه" أو "لا شيء مما ورد أعلاه" أو "أكثر من شيء مما ورد أعلاه".

١٢- لا تستخدم الأسئلة المتداخلة (التي تقود الإجابة على واحد منها إلى الإجابة على الثاني).

١٣- رتب الخيارات على صورة فقرات (تراكمات) وليس على هيئة ازدواجات الواحد خلف الآخر).

١٤- استخدم الأعداد لترقيم الأسئلة والأحرف لترقيم الخيارات.

١٥- تأكد من مراجعة وتدقيق الفقرات من قبل مصمم اختبارات ذو خبرة واسعة قبل استخدامها.

الصندوق ٣-٢

أمثلة عن أسئلة الاختيارات المتعددة المعقدة

١- التصنيف. يطلب من الممتحن أن يصنف الأشياء أو الأشخاص أو الحالات إلى عدد من الفئات التي تم ذكرها في جذر السؤال. مثال:

يمكن وصف جان بياجيه على أنه عالم نفس ------------------.

أ. سريري ب. سايكومتري ج.تطوري د. اجتماعي

٢- حالات إذا - فإن. يحدد الممتحن النتائج الصحيحة لحالة أو عدد من الحالات الموجودة. مثال:

إذا ازداد التباين الحقيقي للاختبار وبقي تباين الخطأ على ما هو عليه، فأي من الأمور التالية سوف

يحدث:

أ- سيزداد ثبات الاختبار.

ب- سيقل ثبات الاختبار.

ج- سينخفض التباين الملاحظ.

د- لن يطرأ أي تغير على الثبات أو التباين الملاحظ.

٣- الشروط المتعددة. يستخدم الممتحن الشروط الموجودة في جذر السؤال (اثنان أو أكثر) ليتوصل إلى استنتاج معين. مثال:

إذا كانت درجة ماري الأولية (Row score) في اختبار ما ٦٠ وكان الوسط الحسابي للاختبار ٥٩ والإنحراف المعياري ٢، فما هو معدل درجات ماري z:

أ. ٢,٠٠- ب. ٠.٥٠- ج. ٠.٥٠- د. ٢,٠٠

٤- أسئلة الصواب والخطأ المتعددة الاختيارات. يقرر الممتحن صحة شرط واحد أو شرطين أو أكثر أو عدم صحة أي من الشروط المذكورة في جذر السؤال. مثال:

هل من الصحيح أن ١) ألفرد بينيه كان أب اختبارات الذكاء وأنه ٢) نشر أول اختبارات الذكاء التي وضعها في عام ١٩١٦؟

أ. ١ و ٢ صحيحان ب. ١ صحيح لكن ٢ غير صحيح

ج. ٢ صحيح ولكن ١ غير صحيح د. ١ و ٢ خطأ

٥- الإختلاف. يشير الممتحن إلى الخيار الذي لا ينتمي لباقي المجموعة. مثال:

أي من الأسماء التالية هو المختلف؟

أ. ألفرد أدلر. ب. سيغموند فرويد. ج. كارل يونغ. د. كارل روجرز

٦- العلاقات والروابط. يحدد الممتحن العلاقة بين المفاهيم ١ و٢ ثم يختار من بين الخيارات (أ، ب، ج، د... إلخ) المفهوم الذي يرتبط مع ٢ بعلاقة مشابهة للعلاقة بين ١ و,٢ مثال:

الوسط الحسابي بالنسبة للانحراف المعياري هو كالوسيط بالنسبة إلى:

أ. الانحراف الوسطي. ب. المدى الشامل.

ج. شبه المدى الربيعي Semi - inter quartile range. د. التباين

وقد أدت صعوبة وضع أربعة أو خمسة اختيارات متساوية الطول وقادرة على جذب الممتحن الذي يجهل جواب السؤال إلى اتجاه واضعي الاختبارات إلى النموذج ذو الخيارين. ويمثل برنامجا F-2 وF-5 الموجودان على القرص المرن المرفق بهذا الكتاب اختبارات تحتمل فقراتها خيارين فقط. يضم القسم الأول اختباراً في تهجئة ٢٨ كلمة غالباً ما يقع الناس في الخطأ عند تهجئتها ويضم القسم الثاني اختباراً في تهجئة ٨ أنواع من الجموع (المفردات في حالة الجمع) بينما يحتوي القسم الثالث على اختبار ٢٠ كلمة لها معانٍ مختلفة. وتوفر هذه البرامج الدرجات الخام والنسبة المئوية للإجابات الصحيحة للاختبار بكامله ولكل قسم على حدة. كما يُستخدم برنامج F-5 لإجراء وتصحيح اختبار مؤلف من ٥٥ فقرة عن المصطلحات المستخدمة في المقررات التمهيدية لمادة علم النفس. يتيح هذا البرنامج للممتحنين فرصة معرفة درجاتهم أثناء الاختبار ويسمح لهم بتغيير إجاباتهم ومن ثم مقارنة إجاباتهم مع الإجابات الصحيحة بعد انتهائهم من الاختبار. ويوفر البرنامج معلومات عن عدد الإجابات الصحيحة

ومعدل الدرجات المصحح من أثر التخمين والنسب المئوية للدرجات النهائية المحتملة.

ترتيب فقرات الاختبار

يعد ترتيب فقرات الاختبار التي تنتمي إلى نوع واحد والتي تتناول موضوعاً واحداً إلى جانب بعضها أمراً من مقتضيات السهولة وحسن التقديم، وعليه ترتب الأسئلة المتعددة الخيارات في قسم خاص وكذلك أسئلة الصح أو الخطأ وهكذا دواليك. وغالباً ما توضع الأسئلة المقالية التي تستغرق الإجابة عليها وقتاً طويلاً والتي تحتاج إلى طريقة تفكير مختلفة من قبل الممتحن في نهاية الاختبار، إلا أنه لا توجد قاعدة ثابتة لهذا الترتيب. وإن عدم تشتت الممتحن وتقليص عدد التحولات الإدراكية التي تطرأ عليه خلال فترة الاختبار هي الأهداف الرئيسة لتقسيم الاختبار إلى أقسام يضم كل منها أسئلة متشابهة. كما ترتب الفقرات الاختبارية عادةً من الأسهل إلى الأصعب لسبب بسيط وهو أن الأسئلة السهلة ترفع من مستوى ثقة الممتحن بنفسه وتشجعه على محاولة الإجابة على الأسئلة الأكثر صعوبة. إلا أن الأبحاث في هذا المجال أثبتت أن ترتيب فقرات الاختبار من الأسهل إلى الأصعب ليس له أي تأثير على الدرجات النهائية (أليسون، ١٩٨٤؛ كليمكو، ١٩٨٤) باستثناء نوعين من الاختبارات وهما الاختبارات الصعبة جداً واختبارات السرعة التي أثبت فيها أن ترتيب الأسئلة من الأسهل إلى الأصعب ساهم في رفع الدرجات بعض الشيء (غرين، ١٩٨٤؛ بلايك وأنسورج وباركر ولاوري، ١٩٨٢؛ سافيتز، ١٩٨٥).

وترتب فقرات بعض اختبارات الذكاء الجماعية على شكل مجموعة من الاختبارات الفرعية المنفصلة أو في اختبارات أخرى فإنها ترتب على شكل كتيبات عادية أو لولبية ترتب فيها الفقرات الاختبارية التي تقيس العمليات المختلفة التي تساهم في تكوين القدرة العقلية للأفراد من الأسهل إلى الأصعب.

وفي هذا النوع من الكتيبات تظهر الأسئلة التي تتناول موضوعاً معيناً بشكلٍ متناوب مع أسئلة من أنواع أخرى ولكن على نفس القدر من الصعوبة وترتب لولبياً من الأسهل إلى الأصعب.

تعليمات الاختبار

تكتب تعليمات الاختبارات الموضوعية أو المقالية التي تجرى في الوقت ذاته لمجموعة من الأشخاص على ورقة توضع في بداية الاختبار بينما توضع التعليمات الخاصة بأقسام الاختبارات المتعددة الفروع في بداية كل قسم. ويفضل طباعة هذه التعليمات بالحروف السوداء العريضة لتقليل احتمالية عدم انتباه الممتحنين إليها. وفي حال كون تعليمات الاختبار غير مألوفة أو معقدة، يفضل أن يقرأها المراقب أو القائم على الاختبار بصوتٍ عالٍ. أما في الاختبارات الفردية التي تسمح بالتواصل بين الممتحن والقائم على الاختبار، يعطي هذا الأخير التعليمات للممتحن شفهياً. ويشترط في تعليمات الاختبار(مقروؤة كانت أم مطبوعة أم الاثنين معاً) أن تُطلع الممتحن على الأمور التالية:

- الهدف العام من الاختبار أو الفقرة الاختبارية.
- الوقت الزمني المتاح للممتحن.
- كيفية وضع الدرجات.
- ما إذا كان التخمين في الإجابة أمراً ينصح به (في حالة الاختبارات الموضوعية فقط).

تمثل التعليمات أدناه نموذجاً عن تعليمات الاختبارات:

اكتب اسمك في الزاوية اليمنى أعلى الصفحة ولا تكتب على ورقة الأسئلة. يهدف هذا الاختبار إلى قياس درجة معرفتك في مادة الإحصاء للمبتدئين. تتوفر أوراق مسودة يمكن استخدامها لإجراء العمليات الحسابية. يتألف الاختبار من ٦٠ سؤالاً ولديك ٩٠ دقيقة لإتمام الاختبار. حدد إجابتك على كل سؤال بملء

الفراغ المناسب على ورقة الإجابة تحت الحرف الذي يمثل الإجابة الصحيحة. سيكون معدل درجاتك الأولي هو مجموع إجاباتك الصحيحة. وعلى الرغم من أن التخمين الأهوج لا يفيد كثيراً إلا أنه من الحكمة أن تحاول استبعاد أي من الخيارات التي تعلم أنها خاطئة قبل اختيار الإجابة. يسمح لك وقت الاختبار بالإجابة على جميع الأسئلة ومراجعة إجاباتك. في حال انهاء الاختبار قبل انتهاء الوقت، الرجاء الجلوس بهدوء لحين انتهاء الآخرين من الإجابة.

يتوجب على القائم على الاختبار قراءة التعليمات بوضوح وبطء وبالشكل الذي تظهر عليه وبعد الانتهاء من القراءة يعطى الممتحنون الفرصة لطرح الأسئلة التي يتعين على القائم على الاختبار الإجابة عليها (بشكل هادئ ومفيد) وإن كانت سخيفة أو مكررة.

أوراق الإجابة

لا ينصح باستخدام الاختبارات التي تتطلب أوراق إجابة منفصلة مع مجموعات الأطفال دون سن التاسعة أو العاشرة (أيراسيان وتراسي، ١٩٩٤) إذ من الأفضل الطلب من الأطفال دون هذا العمر كتابة الإجابات في هامش ورقة الأسئلة. أما في سنوات الدراسية الابتدائية العليا يسهّل استخدام أوراق إجابة منفصلة عمليتي وضع الدرجات وتحليل الإجابات. وفي هذا النوع من الاختبارات التي تصحح عن طريق الآلات، يطلب من الممتحنين ملء الدائرة المرقمة (بالأحرف أو الأرقام) والموجودة إلى جانب السؤال.

إجراء الاختبار

تتأثر درجات الاختبارات النفسية ليس بمحتوى الاختبارات فحسب بل بالطريقة التي يتم فيها إجراء الاختبار. وبالرغم من الجهود المضنية التي تبذل لوضع اختبارات مناسبة من حيث الشكل والمضمون إلا أن هذا الأمر لا يضمن

وجود الدافع لدى جميع الممتحنين للإجابة على أسئلة الاختبار بشكل صادق وملتزم كما لا يضمن فهم هؤلاء الممتحنين لما هو مطلوب منهم. علاوة على ذلك، من السهل أن يصاب الممتحنون بالإرهاق أو أن يتشتت انتباههم أو، لأسباب أخرى، يمكن أن يعاني الممتحنون من فقدان التركيز أو فقدان الرغبة في إتمام الاختبار. وتساهم هذه العوامل مجتمعة في التأثير على معدلات درجات هؤلاء الأفراد مما يعني أن الاختبار فشل في رسم صورة دقيقة عن إمكانياتهم.

وتتنوع الإجراءات التي تُتخذ لضمان أن نتائج اختبارات الاستعداد تقيس بشكل فعال ودقيق المتغيرات التي أعدت لقياسها بتنوع الاختبارات بحد ذاتها. إذ تتطلب الاختبارات الفردية درجة أعلى من الاهتمام من قبل القائم على الاختبار مما تتطلبه الاختبارات الجماعية. وتختلف الإجراءات المرتبطة باختبارات الاستعداد عن تلك المرتبطة باختبارات الشخصية. وعلى إختلاف الاختبارات (فردية أم جماعية محددة بوقت معين أم وقتها مفتوح)، وبغض النظر عن الوقت الذي يجرى فيه الاختبار وعن الخلفية الحضارية أو العرقية للممتحن، يتوجب على القائم على الاختبار أن يكون على درجة عالية من الاستعداد. ويجب، لضمان أن تكون نتائج الاختبارات مقاييس صادقة لما يهدف الاختبار إلى قياسه، إعداد وإجراء الاختبار بشكل حذر والتأكد من أن القائم على الاختبار لديه الخبرة والتدريب اللازمين لإجراء الاختبار وأن الجو في قاعة الاختبار مناسب ومريح.

يتطلب إجراء اختبارات الذكاء واختبارات الشخصية الفردية أن يكون القائم على الاختبار لديه المؤهلات المطلوبة والتي تكون عادة أكثر صرامة من المؤهلات المطلوبة لإجراء الاختبارات الجماعية. ومن الضروري أن يكون القائمون على الاختبار على دراية تامة بحدود إمكاناتهم وأن يطلبوا المساعدة من الأفراد المؤهلين إذا نقصتهم الخبرة أو المعرفة. وفي حال اختبار القاصرين يجب

الحصول على موافقة ولي الأمر أو الوصي القانوني قبل إجراء الاختبار. ويفيد توقيع إستمارة مثل تلك الموجودة في الصندوق ١-١ في حل هذه المشكلة. وهنالك إستثناءات لهذه القاعدة إذ لا يطلب توقيع إستمارة الموافقة على الخضوع للاختبار في حال كون الاختبار ضرورة قانونية منصوص عليها أو في حال كون الاختبار جزءاً من النشاطات المدرسية الاعتيادية أو أن تكون الموافقة على الخضوع للاختبار متضمنة في عقد ما كطلب التوظيف أو طلب الإنتساب إلى كلية أو مؤسسة تعليمية ما.

وعلى الرغم من أن الأشخاص غالباً ما يرقون إلى مستوى المناسبة ويكون أداؤهم حسناً حتى تحت أحلك الظروف فإن الإضاءة وتوزيع المقاعد والتهوية والحرارة ومستوى الضجيج كلها تساعد على ارتفاع مستويات الأداء في الاختبارات. ويقوم القائم على الاختبار بالتعرف على قاعة الاختبار (المقاعد، الإضاءة، موقع النوافذ ومظّمات الحرارة... إلخ) والتعرف على الحمامات والمخارج القريبة ثم يجري تعديلات إذا لزم الأمر. ويفيد وضع لوحة مكتوب عليها "امتحانات! يرجى الهدوء" في تخفيف المقاطعات والضجيج. وفي الاختبارات المدرسية ينصح بعدم إجراء الاختبارات بعد الغداء أو قبل العطلة أو غيرها من الأوقات التي يكون الطلاب فيها متعبين أو منفعلين.

ضمان سلامة الاختبارات والغش

بالإضافة إلى حاجتها لموظفين خبراء، يتطلب إجراء الاختبارات التي يسهل الغش فيها مثل اختبار التقييم المدرسي (SAT) واختبارسجل الخريجين (GRE) إلى رقابة مشددة أثناء فترة الاختبار. ترقم أوراق الاختبار بالتسلسل مما يسهل عملية إحصاء هذه الأوراق قبل وبعد الاختبار كما يطلب من الممتحنين إبراز هوياتهم الشخصية أو أي نوع آخر من بطاقات الهوية ولا يسمح لهم بمغادرة قاعة الاختبار إلا بعد انتهاء الاختبار.

وبغض النظر عن نوعية الاختبار، يبقى خطر الغش واقعاً تعاني منه عملية الاختبار، وغالباً ما يحصل الغش عندما يتمكن أفراد غير مخولين من الحصول على نسخ من أوراق الأسئلة أو غيرها من المعلومات عن فقرات الاختبار. ومن أكثر أشكال الغش شيوعاً إحضار أوراق صغيرة تحتوي على إجابات إلى قاعة الامتحانات أو نسخ الإجابات من ممتحنين آخرين. ومن الجدير بالذكر أنه لا توجد طريقة أكيدة لمعالجة هذه المشكلة ولكن ينصح بوضع مقاعد الممتحنين على مسافات بعيدة (وترك مقعد واحد فارغ على الأقل) والطلب من الممتحنين إفراغ محتويات جيوبهم ومقاعدهم وتكثيف عدد المراقبين في قاعة الامتحان والتأكد من أوراق الأسئلة قبل وبعد الاختبار. كما يساعد برنامج B-7 الموجود على القرص المرافق لهذا الكتاب في التعرف على عدد من حالات الغش في الاختبارات.

يمكن أن تحدث حالات الغش قبل أو بعد أو أثناء الاختبار فعلى سبيل المثال، إكتشفت منذ عدة سنوات حالة غش واسعة النطاق في كاليفورنيا إذ كان المدرسون في بعض مدارس الولاية يمسحون الإجابات الخاطئة التي اختاراها الطلاب في اختبار برنامج كاليفورنيا للتحصيل (CAP) ويضعون الإجابات الصحيحة مكانها وذلك بهدف تحسين مظهر مدارسهم وطلابهم. كما يمكن أن تحدث حالات غش بعد انتهاء الاختبار إذ يمكن إضافة نقاط إلى معدلات درجات طالب أو موظف معين بهدف جعل هذا الفرد يبدو أكثر كفاءة لأي سبب من الأسباب.

تعليمات الاختبار وإجراءاته

من المهم التقيد بتعليمات الاختبارات القياسية وخاصةً في الحالات التي سيتم تحويل درجات الاختبار فيها إلى معايير عامة مستمدة من مجموعة الأفراد الذين تم اختبارهم وفق هذه التعليمات. إلا أنه، وفي حالات الاختبارات

السريرية، لا يمكن دائماً التقيد الدقيق بتعليمات الاختبار ففي حالة اختبار شخص معاق جسدياً أو اختبار حدود الإمكانيات الجسدية أو العقلية لفردٍ ما يفضل عدم التقيد بالتعليمات النموذجية.

يتعين على القائم على الاختبار في حالة الاختبارات الجماعية والفردية التعريف بنفسه ثم التحدث مع الممتحنين للتعرف عليهم بشكل سريع. وفي حالة الاختبارات الفردية تتطلب عملية التعرف والتواصل هذه جهداً أكبر ووقتاً أطول إذ يتعين على القائم على الاختبار الوصول إلى درجة جيدة من التواصل مع الممتحن بهدف إعطاء الممتحن شعوراً بالطمانينة مما يمكن أن يزيل مخاوف الممتحن ويساعده على تحقيق مستوى أداء أعلى في الاختبار. وغالباً ما تساعد خبرة القائم على الاختبار في التعامل مع أفراد من نفس العمر والجنس والعرق والخلفية الحضارية واللغوية الذي ينتمي إليه الممتحنون على مساعدة الممتحنين على الوصول إلى مستوى أدائهم الأمثل. ويتطلب اختبار أفراد متقدمين في العمر أو أطفال صغار أو أفراد لديهم إعاقات معينة أو أفراد ينتمون إلى حضارة مختلفة الخروج عن إجراءات الاختبار وتعليماته النموذجية، فيمكن مثلاً تغيير صياغة الأسئلة لجعلها أكثر بساطة أو زيادةوقت الاختبار أو تقصيره أو تشجيع الممتحنين بشكل دوري ومتواتر أو توفير أدوات مساعدة على التحرك أو الإحساس. ويجب مراعاة في حال تعديل إجراءات الاختبار لمجموعة معينة من الأفراد أنه لا يمكن مقارنة أداء هذه المجموعةمع أداء عينة القياس. ومن الحكمة أحياناً عند اختبار الأطفال المعاقين أن يعطى هؤلاء الأطفال اختبارات أدنى من مستوى عمرهم أو سنتهم الدراسية (وهذا ما يسمى الاختبار الخارج عن المستوى). وتتوفر في هذه الحالات معايير خاصة للاختبار الخارج عن المستوى تقيم أداء الممتحنين في الاختبارات القياسية.

البراعة في الاختبارات

يُعتبر بعض الأفراد أقدر من غيرهم على تحقيق أداء جيد في الاختبارات ويعزى هذا إلى عدد من العوامل كمستوى ذكاء أعلى أو مستوى قلق أدنى أو بسبب كون الطالب أكثر جاهزية أو أبرع في الاختبارات لسبب مجهول. وغالباً ما يرتفع مستوى القلق لدى الأفراد في الاختبارات المحدودة الوقت، ومما لا شك فيه أن هذا القلق يزيد عندما يكون الممتحن على درجة غير مقبولة من الجاهزية للاختبار. وعلى الرغم من أن القلق وعدم الجاهزية يؤثران على معدلات درجات الاختبارات، يلعب عامل يصعب تعريفه هو عامل "البراعة في الاختبارات" دوراً في التأثير على الدرجات. ويمكن لشخص بارع في الاختبارات وجاهل بمحتوى الاختبار أن يرفع مستوى أدائه عن طريق الاهتمام بإجاباته (إيفانز، ١٩٨٤). ويعد بعض الممتحَنون أكثر قدرة على التعرف على الإجابات الخاطئة في الأسئلة متعددة الخيارات حيث يمكنهم ملاحظة أن الإجابات الخاطئة غالباً ما تصاغ بطريقة مختلفة عن الإجابات الصحيحة (تكون الإجابات الخاطئة عامة جداً أو محددة جداً أو تحتوي على أخطاء نحوية). ويلاحظ هؤلاء الأفراد كذلك كون الإجابات الصحيحة دقيقة التعبير وتحتوي على دلائل نحوية أو أنها مرتبطة بإجابات أو فقرات اختبارية أخرى. وتتوفر أدلة على كون فقرات الاختبار اللغوية أكثر تأثراً بعامل "البراعة في الاختبارات" وعلى كون الذكور أكثر "براعة" من الإناث (برستون، ١٩٦٤؛ راولي، ١٩٧٤). هذا ويمكن تدريب الأفراد لزيادة براعتهم في الاختبارات (وهذا ما يوضحه الصندوق ٣-٣).

الاختبارات المحوسبة

توضح برامج الكمبيوتر B-2 وبرامج الفئات F وG وH الموجودة على القرص المرافق لهذا الكتاب سهولة إجراء ووضع درجات الاختبارات الموضوعية بوساطة الحاسب الآلي. وتعد برامج الاختبارات التكيفية المحوسبة

Computer - based adaptive testing (CAT) والتي تحدد الأسئلة اللاحقة فيها بأداء الممتحَن في الأسئلة السابقة نوعاً مطوراً من برامج الاختبارات المحوسبة السهلة. ويقوم برنامج الكمبيوتر باختيار السؤال اللاحق لممتحن ما بناءاً على إجاباته على الأسئلة السابقة، ويتم الاختيار من مجموعات من الأسئلة تمت معايرتها تبعاً لمستويات القدرة المختلفة وبناءً على نموذج الاستجابة للأسئلة IRT. يبدأ الممتحن بسؤال يعد مستوى صعوبته متوسطاً ومن الممكن أن يكون هذا السؤال مبنياً على عمر الممتحن أو درجاته في اختبارات سابقة أو غيرها من المعلومات الإضافية، وبعد فراغ الممتحن من الإجابة على السؤال الأول يحدد الكمبيوتر مستوى السؤال التالي الذي يكون من مستوى قدرة أعلى في حال إعطاء الممتحن إجابة صحيحة على السؤال أو من مستوى قدرة أدنى إذا أعطى الممتحن إجابة خاطئة، وتكرر هذه العملية طوال الاختبار.

يسمح هذا النمط من الاختبارات بالحصول على تقدير دقيق لقدرة الممتحن من خلال عدد من الأسئلة يشكل ٥٠٪ من عدد الأسئلة الموجودة في كتيبات الاختبار التقليدية.

الصندوق ٣-٣

اقتراحات لزيادة "البراعة" في الاختبارات

قبل الاختبار

١- اطلب من المدرس إطلاعك على أسئلة السنوات السابقة من نفس الاختبار والتي يُسمح لك الإطلاع عليها.

٢- إسأل الطلاب عن أنواع الاختبارات التي يأتي بها مدرس المقرر.

٣- لا تؤخر التحضير للاختبار إلى اليوم الذي يسبق الاختبار في حال كون يوم الاختبار قد أعلن عنه مسبقاً.

٤- ادرس بشكل يتناسب مع نوع الاختبار المعلن عنه (أسئلة متعددة الاختيارات، صواب أو خطأ، مقالات).

٥- في حال عدم الإعلان عن نوع محدد من الاختبارات ينصح بالتحضير للاختبارات المقالية التي تعتمد على الذاكرة.

٦- لا تحول عملية التحضير للاختبار إلى مناسبة اجتماعية إذ من الأفضل لك أن تنعزل وتدرس بمفردك.

٧- لا تجلس بشكل مريح جداً (كالاستلقاء على السرير) إذ قد يسيء جسمك فهم هذه الراحة ومن الجائز أن تشعر بالنعاس.

٨- حاول أن تُعد أسئلة اختبار من المعلومات التي تدرسها وحاول أن تكون هذه الأسئلة من نوع الخيارات المتعددة إذا كان الاختبار يعتمد على هذا النوع من الأسئلة، أو أسئلة مقالية إذا كان الاختبار مقالياً.

٩- حاول أن تطبق طريقة الدراسة المعروفة بـ Q3R (خذ فكرة عامة عن المادة أو المقرر، ضع أسئلة، اقرأ، استذكر، راجع). اقرأ أولاً لتكوّن فكرة عامة عن المادة العلمية ثم إسأل نفسك بعض الأسئلة حول هذه المادة ثم إقرأ المادة بتمعن يمكّنك من تذكر المعلومات ثم كرر المعلومات لتثبت في ذاكرتك وأخيراً راجع المعلومات قبل الاختبار.

أثناء الاختبار

١- اقرأ تعليمات الاختبار بتمعن قبل البدأ بالاختبار وفي حال وجود أي استفسارات حول أشياء لم تذكرها التعليمات كالزمن المتاح لإنهاء الاختبار أو أوزان فقرات الاختبار المختلفة أو ما إذا كان الممتحن يفقد درجات في حال التخمين الخطأ للإجابة أو غيرها، لا تتردد في سؤال المراقب أو القائم على الاختبارعنها.

٢- فكر ملياً بالإجابة أو استعمل المسودة قبل الإجابة على الأسئلة المقالية.

٣- وزع وقتك بالتساوي على الأسئلة ففي حالة اختبارات الإختيار من متعدد يجب أن تجيب على عدد مناسب من الفقرات يعادل الزمن الكلي للاختبار مقسماً على عدد الأسئلة.

٤- لا تترك أي سؤال دون إجابة إذا كان بإمكانك إستبعاد أي من الخيارات حتى في حال فقدان الدرجات في حالة التخمين الخطأ للإجابة.

٥- إترك الأسئلة الصعب التي لم تستطع الإجابة عليها وعد إليها بعد فراغك من الإجابة على الأسئلة الباقية. حاول أن لا تصاب بالذعر بسبب الأسئلة التي لا تعرف إجاباتها. ضع إشارة إلى جانب هذه الأسئلة وحاول أن تجيب عليها وإن لم تستطع فخمن عن علم وتفكير.

٦- في حال توفر الوقت لديك لمراجعة إجاباتك إفعل هذا ولا تكن في عجلة لتسليم ورقة الاختبار ومغادرة القاعة.

تتوفر عدد من الاختبارات تعتمد على مبدأ مشابه للاختبارات التكيفية كالاختبارات المتعددة المستويات والاختبارات ذات المستويات المرنة. وتعد الطبعة الرابعة من اختبار ستانفورد - بينيه واحدة من هذه الاختبارات المتعددة المستويات التي تعتمد على إجراء اختبار قصير لجميع المتحنين يطلق عليه اسم الاختبار "محدد المسار" وهو اختبار تحدد معدلات درجات المتحن فيه مستوى الاختبار الذي سيعطى للممتحن لاحقاً. وترتب الأسئلة في الاختبارات ذات المستويات المرنة تبعاً لمستوى الصعوبة من الأسهل إلى الأصعب ويطلب من جميع المتحنين البدء بالأسئلة التي تقع في المنتصف ثم تؤهل الإجابات الصحيحة المتحن للإجابة على الأسئلة الأكثر صعوبة بينما تقود الإجابات الخاطئة المتحن إلى الإجابة على الأسئلة الأكثر سهولة وتكرر هذه العملية طوال الاختبار (غرين، ١٩٩٢).

تتمتع الاختبارات التكيفية بمزايا عديدة ككونها أكثر دقة من أشكال الاختبار التقليدية وكونها تسمح بخلط احتمالات الإجابة لمنع تكرار الخيارات ذاتها كما تسمح باختبار كل ممتحن بناءاً على مستوى قدرته وبأن يحس الممتحن بالتحدي طوال الوقت. إلا أن أحد المشاكل المرتبطة بهذا النوع من الاختبارات هي عدم قدرة الممتحن على ترك سؤال والعودة إليه لاحقاً أو تغيير إجابته على السؤال، كما يتوجب على واضع الاختبار التأكد من وقوع السؤال برمته على صفحة واحدة تظهر على شاشة الكمبيوتر بشكل كامل. كما تعاني الاختبارات التكيفية من المشاكل التي تعاني منها نظرية تحليل الفقرات الاختبارية التي تعد أساس الاختبارات التكيفية. إلا أنه، وعلى كل حال، تعتبر الاختبارات التكيفية (CAT) اختبارات يستجيب إليها الممتحنون بشكل جيد (غرين، ١٩٩٢).

تصحيح الاختبارات وتفسير النتائج

يجب إتخاذ القرارات المتعلقة بتصحيح الاختبارات (إجراءات التصحيح وقواعده ومفتاح الإجابات) قبل إجراء الاختبار وخلال مرحلة التصميم. كما يجب تحديد الطريقة التي سيتم من خلالها التعبير عن نتائج الاختبارات عددياً، ويجب اعتماد طريقة لوضع الدرجات إما باليد أو عن طريق الحاسب. ومن الجدير بالذكر أن استخدام ورقة منفصلة للإجابة لا تعني بالضرورة استخدام الحاسب لوضع الدرجات وإن كان هذا الاستخدام يسهل من عملية التصحيح المحوسبة. وسواء كان التصحيح باليد أو بالحاسب الآلي، هل سيتم التصحيح على هيئة سلسلة من الأجزاء أو على الاختبار ككل أو كليهما؟ ومن الضروري إتخاذ قرار بشأن عدد الدرجات التي تعطى للإجابات الصحيحة ويتم الاختيار عادةً بين طريقة التصحيح البسيطة التي تمنح درجة واحدة للسؤال الصحيح أو طريقة التصحيح الأكثر تعقيداً التي تعطي أوزاناً مختلفة للإجابات المختلفة. كما يجب، في الاختبارات الموضوعية (أسئلة الاختيار المتعدد وأسئلة الصواب والخطأ)، أخذ مبدأ "التصحيح من أثر التخمين" بعين الاعتبار وهو مبدأ يتم على أساسه حذف

عدد من الدرجات الأولية للممتحنين الذين خمنوا خطأً عدداً من إجاباتهم. وفي الأسئلة المقالية يجب تحديد ما إذا كان يحب تقييم الإجابات بناءاً على عاملين منفصلين كالمحتوى والشكل (القواعد والترتيب... إلخ)، وإن كان الأمر كذلك هل يجب الجمع بين هذين العاملين بشكلٍ أو بآخر. والسؤال الأخير الذي يجدر طرحه هنا هو: هل سيتم تحويل النتائج الأولية للاختبارات إلى علامات نهائية (A، *A... إلخ) وهل يجب إرفاق ملاحظات مكتوبة مع ورقة النتائج؟

اختبارات الخيارات المتعددة والخطأ والصواب

تعتمد أبسط طرق وضع درجات الاختبارات المؤلفة من أسئلة الصواب والخطأ أو أسئلة الخيارات المتعددة على إعطاء درجة واحدة للإجابة الصحيحة ودرجة الصفر للإجابة الخاطئة مما يعني أن الحد الأدنى من الدرجات الممكن الحصول عليها هو صفر والحد الأقصى هو العدد المساوي لعدد فقرات الاختبار. وتعد طريقة إعطاء درجات أو أوزان مختلفة للأسئلة الأكثر أو الأقل صعوبة أو لخيارات مختلفة للسؤال ذاته طريقةً منطقية وعادلة إلا أنه وفي الاختبارات التي تتجاوز فقراتها العشرين لا يكون تأثير مثل هذه الاختلافات في الأوزان محسوساً على تمييز الفقرات الاختبارية أو على ثبات الاختبار. كما لم يتم إثبات أن ما يسمى بوزن الثقة Confidence weighting (حيث توزع الدرجات على أساس نقاط الثقة التي يعطيها الممتحن لإجابته والتي تعبر عن درجة ثقته في صحة إجابته) هو أكثر دقة من المقاييس التقليدية التي تعتمد على مبدأ صفر - ١ في وضع الدرجات. ومن أكثر طرق وضع الأوزان شيوعاً الطريقة المستخدمة في نظرية الاستجابة للأسئلة والتي تعتمد على النظر إلى درجات الاختبارات على أنها تعبير عن درجات تقديرية تقع ضمن مجال القدرة الكامنة. كما تقوم بعض الاختبارات مثل اختبارات الذكاء الفردية (كمقياس ويكسلر لذكاء البالغين) بإعطاء درجات مختلفة (٠ أو ١ أو ٢) للإجابات وفقاً لدقتها أو شموليتها أو لسرعة التوصل إليها.

ويرتبط إجراء "التصحيح من أثر التخمين" باختلاف أوزان الفقرات المختلفة. ويطبق هذا التصحيح على عدد الإجابات الصحيحة في الاختبارات الموضوعية باستخدام المعادلة التالية:

$$s = r - w/(k-1) \qquad (٢-١)$$

وتستند هذه المعادلة إلى الاعتقاد البسيط بأن الممتحَنين الذين لا يعرفون الإجابة على سؤال ما سوف يقومون باختيار واحد من الخيارات عشوائياً. وفي المعادلة تعبر s عن درجات الاختبار بعد تصحيح التخمين، و r عدد الأسئلة التي أجاب عليها الممتحن بشكل صحيح، و w عدد الإجابات الخاطئة، و k عدد البدائل لكل سؤال. وقد تم اقتراح معادلات أخرى لتصحيح التخمين بسبب الاعتقاد بأنه لا يجب أن يلعب التخمين العشوائي دوراً في رفع درجات الأفراد. فعلى سبيل المثال تبدو معادلة ليتل (١٩٦٢) على النحو التالي: $s = r - w/[2(k-1)]$، ومعادلة ريد (١٩٧٧) $s = (n+r)(kr - n)/[2r(k-1)]$، حيث n هي العدد الإجمالي لفقرات الاختبار. لا تتضمن معادلة ريد عدد الإجابات الخطأ ويمكن أن ينتج عنها معدل درجات أعلى من ذلك الناتج من المعادلة التقليدية التي فيها $s = r$. ويساعد برنامج B-3 الموجود على القرص المرافق لهذا الكتاب في حساب التصحيح من أثر التخمين وفقاً للمعادلة المتعارف عليها ومعادلات ريد وليتل.

وتم استكشاف طريقة أخرى للأسئلة الموضوعية وهي إجراء وضع الدرجات عن طريق الاستبعاد وينطوي التغيير في هذا الأسلوب على تعيين درجتين؛ واحدة تحدد للخيارات الصحيحة وأخرى للخيارات الخاطئة، ثم تحسب الدرجة النهائية باعتبار وزنين متساويين من الدرجتين (إدوارد وهاريس، ١٩٨٣). وعلى الرغم من أن كوليت (١٩٧١) وجد أن طريقة وضع الدرجات عن طريق الاستبعاد هي أكثر

صدق وثبات من طريقة التصحيح من أثر التخمين أو الأوزان المختلفة لفقرات الاختبار، وجد هاكستيان وكانسوب (١٩٧٥) أن اختلاف الأوزان المعطاة للإجابة على الفقرات الاختبارية لا يزيد من صدق أو ثبات الاختبار.

أسئلة إعادة الترتيب

توضع درجات اختبارات المطابقة عن طريق إعطاء درجة واحدة لكل مطابقة صحيحة ودرجة الصفر للمطابقات الخاطئة. أما في حالة فقرات الاختبار التي تتطلب إعادة ترتيب محتويات السؤال أو تدريجها تسلسلياً فيصعب استخدام هذه الطريقة في وضع الدرجات التي تعتمد على مبدأ كل شيء أو لا شيء (أي أن الإجابة إما خاطئة تماماً أو صحيحة تماماً)، إذ على الرغم من أن عدد العناصرالمطابقة والمراد ترتيبها يمكن أن يكون نفسه في سؤالين مختلفين من الممكن أن يكون عدد العناصر التي وضعها الممتحَن في المكان الخطأ أقرب إلى عدد العناصر كلها. ولهذا السبب ينصح باستخدام المعادلة التالية عندما يطلب من الممتحَنين ترتيب مجموعة من العناصر (الأشياء أو الأحداث) m بناءً على صفة معينة (كالحجم أو العمر أو التاريخ... إلخ):

$$s = m[1 - 2\sum \mid d \mid /(m^2 - j)] \qquad (٢-٣)$$

وفي هذه المعادلة m هو عدد الأشياء التي تم إعادة ترتيبها وd هو الفرق بين الترتيب الصحيح والترتيب الذي اختاره الممتحَن و0 = j إذا كان m عدداً زوجياً و= ١ j إذا كان m عدداً فردياً و s هو معدل درجات الممتحَن للفقرة الاختبارية. يساعد البرنامجان B-4 أو B-5 على استخدام المعادلة ٣-٢ ومعادلة بديلة (راجع آيكين، a١٩٩٧، ص ٥٩)، للحسابات التي تعطي وزناً أكبر للفروقات الأكبر في المراتب (أي تقتطع درجات أكثر من مجموع الممتحن عندما يرتب العناصر المراد ترتيبها بشكل بعيد جداً عن الترتيب الصحيح).

الاختبارات المقالية

تعد الاختبارات المقالية اختبارات صعبة التصحيح وسهلة التصميم على عكس الاختبارات الموضوعية التي يسهل تصحيحها ويصعب تصميمها. وقد جرى مؤخراً العمل على تطوير طريقة لتقييم الإجابات المقالية عن طريق إجراءات كمّية تتعرف على الملامح الرئيسة للإجابات إلا أن هذه الطريقة لا تزال مكلفة وغير موثوقة ولا تسمح، في الوقت الراهن على الأقل، بالاستغناء عن المصحح التقليدي.

تكمن صعوبة تقييم الأسئلة المقالية في كون هذه الأسئلة ذاتية إذ لا يفهم جميع الممتحنين السؤال بالشكل نفسه كما أن لكل ممتحَن أساليبه المختلفة في الكتابة والتعبير عن الأفكار. ولهذا السبب تختلف محتويات الإجابات وأساليب الكتابة إختلافاً شديداً من شخص إلى آخر. حتى عندما تتوجه الجهود لجعل وضع درجات للاختبارات المقالية أكثر موضوعية وذلك بتحديد أوزان معينة لعناصر المقال المختلفة، فإن هذه العملية تظل معقدة ورتيبة وتتطلب وقتاً كبيراً وليست على درجة عالية من الصدق. مع هذا فإن الدرجات التي تستنبط من مثل هذه العملية التحليلية لوضع الدرجات تعد طريقة أكثر ثباتاً من الطريقة الشمولية في وضع الدرجات والتي تنظر إلى المقال بشكل عام بدلاً من النظر إلى عناصره المختلفة. يضع برنامج ستانفورد لتقييم الكتابة (الطبعة الثالثة) درجات لمقالات الطلاب على مقياس مؤلف من ست نقاط بناءً على نظرة شمولية للمقال برمته. كما يحتوي الكتاب أيضاً على نظام وضع الدرجات التحليلي الذي يعتمد على تقسيم ما يُقيِّم المقال على أساسه إلى ست مجموعات هي: الأفكار والتوسع فيها؛ التنظيم والإنسيابية والتناغم؛ اختيار المفردات؛ الجمل والمقاطع؛ النحو والدقة اللغوية؛ وتقنيات المقال.

ويمكن تدريب المصححين أو قارئي الإجابات (كما يسميهم برنامج المجلس الأعلى للقبول الجامعي) لكي يصبحوا أكثر موضوعية وأكثر ثباتاً على مبادئ وضع الدرجات. وقد درجت الجامعات والكليات المختلفة منذ زمنٍ بعيد وقبل أن تظهر الاختبارات ذات أسئلة الخيارات المتعددة على استخدام الأسئلة المقالية لتحديد الطلاب الذين سيتم قبولهم في هذه الجامعة أو الكلية. ولا تزال خدمة الاختبارات التعليمية Educational Testing Services تستعين بعدد كبير من المدرسين والخبراء سنوياً لتقدير الإجابات على الأسئلة المقالية لاختبارات القبول الجامعي المتقدمة التي يخضع لها طلاب المدارس الثانوية في شتى المواد وفي مختلف أنحاء الولايات المتحدة. يجتمع هؤلاء الخبراء كل عام في مدينة برينستون في ولاية نيوجرسي حيث يتم تدريبهم على ما يتعين عليهم البحث عنه في إجابات الطلاب على اختلاف مستويات دقتها، ويشرف على هؤلاء الأفراد "رئيس المصحّحين" الذي يعطي تعليماته بين الحين والآخر لإعادة قياس طريقة التصحيح وإعادة النظر في المقاييس المعتمدة. ويمكن اتباع إجراء مماثل على مستوى المدرسة وبين مدرسي الفصول المختلفين إذ يتم اختيار إجابات للأسئلة المقالية تنتمي إلى عدد من مستويات التحصيل قبل البدء بعملية وضع الدرجات. ويمكن في حال استخدام طريقة وضع الدرجات الشمولية إعطاء درجة عامة للمقال بين ١ و٥ (كما في اختبارات القبول الجامعي المتقدمة المذكورة أعلاه). أما في حال استخدام طريقة وضع الدرجات التحليلية، فإن تقديرات منفصلة تعطى لعناصر مختلفة من الإجابة المراد تقييمها. وفيما يلي مقترحات أخرى يمكن أن ترفع مستوى صدق وثبات الاختبارات المقالية:

١- يجب وضع درجات سؤال ما على أوراق جميع الممتحَنين قبل الإنتقال إلى السؤال التالي.

٢- يجب وضع درجات السؤال ذاته على أوراق جميع الممتحنين خلال المدة الزمنية ذاتها.

٣- يجب تقييم الناحية الشكلية (كالأسلوب وتقنيات الكتابة ونوعيتها) بمعزل عن تقييم المحتوى في الحالات التي يتم تقييم المقال فيها على أساس هذين العاملين.

٤- يزداد ثبات الدرجات عندما يتم حسابها كمعدلات وسطية لدرجات وضعها مصححان أو أكثر.

٥- يجب تصحيح الأخطاء على ورقة الإجابة وكتابة تعليق إلى جانب كل إجابة.

تقييم الدرجات وتفسيرها

لا تعني درجات الاختبارات أي شيء إن لم تتم مقارنتها مع توقعات أو مقاييس أو معايير معينة. ففي حالة الاختبارات ذات المرجعية المحكية تتم مقارنة درجات الأفراد مع المعيار المتفق عليه والذي يحدد مستوى الأداء الفعال ثم يتم تحويل الدرجات الأولية إلى نسب مئوية تحدد موقع الطالب بالنسبة إلى المعيار العام. ويتم عادةً تحديد المعيار العام لاختبار تحصيلي ذو مرجعية محكية من قبل لجنة مكونة من عدد من المدرسين أو الإداريين في المدارس أو خبراء التقييم أو غيرهم من الأفراد المطلعين على التقدم في التعليم وأساليب تقييمه.

وعلى عكس الاختبارات ذات المرجعية المحكية تُقارن معدلات الدرجات في الاختبارات ذات المرجعية المعيارية بعدد من المعايير التي تم التوصل إليها عن طريق دراسة التوزيع التكراري لدرجات تم الحصول عليها من عينة تمثل مجموع الأفراد الذين طور الاختبار من أجلهم. وكما ذكرنا في الفصل الثاني فإن المعدلات أو الدرجات المحوَّلة يمكن أن تكون عمراً أو مكافئات للدرجة أو مراتب مئينية أو درجات قياسية. ومن الجدير بالذكر هنا أنه، وعلى عكس مقاييس الاختبارات ذات المرجعية المحكية، فإن المعايير ليست مقاييس باعتبار كونها أهدافاً أو متطلبات راسخة؛ بل مجرد معدلات وسطية. وتلخيصاً لما سبق يمكن

القول أنه يتم مقارنة درجات اختبار ذو مرجعية محكية بمستوى تم الاتفاق عليه واعتباره مستوى مرغوب فيه للأداء بينما تتم مقارنة درجات اختبار ذو مرجعية معيارية بأداء مجموعة من الأفراد مشابهة للممتحَنين بطرق واضحة.

ولا تعتبر مقاييس الأداء الذاتية التي يضعها المدرسون والمقيمون والمرشدون الاجتماعيون مقاييس ثابتة منقوشة على الحجر بل هي مجرد ملاحظات خلص إليها هؤلاء الأشخاص نتيجة إحتكاكهم بالطلاب. وعليه فإن المقاييس التي يعتمدها المدرسون والأخصائيون الاجتماعيون وأرباب العمل لفهم درجات الأفراد هي مقاييس ضمنية وغير صريحة. وتؤثر عوامل أخرى على هذه المقاييس الذاتية كالتقييم الذاتي وتوقعات المقيِّم أو طموحاته لأداء الممتحنين.

من غير الممكن فهم درجات الاختبارات (في المدارس أو مراكز الإرشاد أو العيادات أو الشركات والمؤسسات) بمعزل عن معلومات متعلقة بخلفية الممتحَنين، إذ يجب الإستعانة بمعلومات عن العمر والمستوى التعليمي والصفات الجسدية والحالة الصحية والخلفية الحضارية والإنجازات السابقة للفرد وغيرها من العوامل المتعلقة بالخبرة والأمور البيولوجية لتفسير درجات الأفراد قبل إتخاذ قرارات تخصهم في مجالات التعليم والتوظيف والعلاج. ومن المهم أن نكون على قناعة تامة بأن نتيجة الاختبار ليست علامة تطبع على جبين الممتحَن لتلازمه طوال حياته. فكل درجة يتم الحصول عليها تحتوي على نسبة من الخطأ التي تختلف مع الوقت، ولذلك تعتبر الدرجات إشارات أو إقتراحات لا يكتمل معناها إلا بالنظر إلى معلومات أخرى عن الممتحَن. وإضافةً إلى ذلك من المفيد مقارنة درجات الأفراد ليس مع الآخرين فحسب بل مع أنفسهم حيث تساعد معلومات عن درجات ممتحَن ما في عدد من الاختبارات الفرعية لسلسلة من اختبارات التحصيل أو الاستعدادات المتعددة على تكوين صورة واضحة عن نقاط ضعف وقوة هذا الممتحن المختلفة في مجال الإدراك.

يصنف البشر الأحداث والأشياء والأفراد وكل ما يشعرون به أو يحسونه بالفطرة ودون أن يجبروا على هذ التصنيف. ويصنف علماء النفس الناس بناءً على درجاتهم في اختبارات القدرة واختبارات الشخصية إلى: أفراد طبيعيين أو معاقين عقلياً أو موهوبين أو مضطربين نفسياً أو مصابين بالعصاب... إلخ. ويصنف العاملون في مجال التعليم الأفراد إلى: طلاب ضعيفين أو بطيئي التعلم أو مصابين بنقص الانتباه وفرط الحركة (ADHD) أو مصابين بمرض صعوبة القراءة والكتابة (ديسلكسيا)... إلخ. كما أن هنالك طلاب ممتازون أو ما يسمى طلاب الدرجة A وطلاب جيدون جداً يحصلون على الدرجة B معظم الوقت وطلاب جيدون يحصلون على الدرجة C، إلا أن هذا لا يعني أن هذه الدرجات موحدة في المقررات المختلفة. وقد تكون درجات الأفراد مستخلصة مباشرةً من الدرجات الأولية للاختبارات ومن الممكن أن يكون المصحح قد أخذ مستوى الأداء العام للفصل وتوقعاته من الطلاب بعين الاعتبار. ويساعد برنامج B-6 الموجود على القرص المرافق للكتاب على إدخال جميع العوامل التي تدخل في حساب الدرجات النهائية من التوزيع التكراري لدرجات الاختبار إلى وسيط مستوى قدرة الفصل التقديري (على مقياس من ١ إلى ١٠٠) إلى الدرجات العظمى والدنيا التي حصل عليها الطلاب إلى الدرجات الأولية لكل ممتحن. وبالرغم من كل هذه التقنيات لا يتم وضع الدرجات النهائية لأي ممتحن دون الرجوع إلى خاصياته الاجتماعية والسياسية والتعليمية.

صياغة نتائج الاختبارات في تقارير رسمية

بالاعتماد على طبيعة الاختبار وعلى كيفية استعمال النتائج، فإن الدرجات سيتم رصدها في دفتر الدرجات الخاص بمدرس المادة، أما في حالة الاختبارات القياسية تُسجل نتائج الاختبارات في سجل تراكمي أو في بطاقة التقرير المدرسي. كما تحفظ درجات المتقدمين إلى مؤسسة تعليمية أو مهنية ما مع

غيرها من أوراق الطلب الرسمية للعودة إليها عند اللزوم. ومن الممكن إرسال نتائج الاختبارات مع غيرها من المعلومات عن شخصٍ ما إلى أفراد مؤهلين يودون الإستفادة من هذه المعلومات.

وعند إحالة قضية طفل أو شخص للدراسة من قبل مدرسة أو مركز إرشاد أو عيادة نفسية، تكون تقارير نتائج الاختبارات وغيرها من إجراءات تقييم قدرات الأفراد تقاريراً مستفيضة. وغالباً ما تتألف هذه التقارير من عدة أقسام تضم معلومات الممتحَن الشخصية (كاسمه وعمره وتاريخ ميلاده وخلفيته التعليمية) واسم القائم على الاختبار ومكان الاختبار وتاريخه ونوع الاختبارات التي أجريت سابقاً ونتائجها بالإضافة إلى الاستنتاجات والتوصيات. ويجب أن تحمل التقارير توقيع المشرف على الاختبار وتاريخ تحرير التقرير. ويجب إعارة قسم الاستنتاجات والتوصيات اهتماماً خاصاً إذ يضم هذا القسم النتيجة النهائية لعمليتي التقييم والتشخيص. ويجب أن تتم صياغة التقريربأسلوب متزن يعبر بوضوح عن الاتجاه النظري للأخصائي الذي درس الحالة والفرضيات النفسية التي بنيت الاستنتاجات النهائية عليها. كما يجب توخي الحذر من حيث التعبير باستخدام لغة الاحتمال بدلاً من لغة التأكيد، وإذا لم يكن بالإمكان الحصول على استنتاجات يمكن تبريرها من خلال النتائج فيجب أن لا تذكر أي استنتاجات.

الخلاصة

تناول هذا الفصل من الكتاب عمليات تصميم ووضع وإجراء الاختبارات ثم تصحيحها وتحليل نتائجها وكتابة تقارير عنها. ويجب قبل البدء بعملية وضع الاختبار تحديد الهدف من الاختبار والمجموعة البشرية المراد اختبارها ووقت الاختبار وطريقة التصحيح وكيفية استخدام نتائج الاختبار. يمكن تصميم اختبارات الاستعداد لتعمل كمقاييس تراكمية أو تكوينية للتحصيل إذ يمكن لهذه

الاختبارات أن تساعد في توزيع الأفراد على الصفوف الدراسية أو الوظائف المختلفة أو أن تساعد في تشخيص ومعالجة بعض المشاكل والاضطرابات أو في تقييم البرامج والمشاريع والمؤسسات. ويتوجب على مستخدم الاختبار التأكد من أن الاختبار مناسب للهدف المراد استخدامه لتحقيقه.

يساعد جدول مواصفات الاختبار أو المخطط التفصيلي للاختبار على تحديد شكل ونوع ومحتوى الاختبار. وعلى الرغم من أن فقرات الاختبار الموضوعية كأسئلة الخيارات المتعددة وأسئلة الصواب والخطأ هي الأسئلة الأكثر شيوعاً في الاختبارات القياسية، تعتبر الاختبارات ذات الأسئلة المقالية مقاييس أفضل للعديد من الأهداف. وإن ترتيب فقرات الاختبار عن طريق وضع الأسئلة المتشابهة في القسم ذاته وترتيب الأسئلة من الأسهل إلى الأصعب قد يكون أمراً مناسباً للممتحَنين والقائمين على الاختبارات، إلا أن مصممي الاختبارات سيحسنون صنعاً عندما يركزون أكثر على كتابة فقرات اختبار جيدة أكثر من تركيزهم على ترتيبها. ويجب أن تكون تعليمات الاختبارات، التي يمكن أن يقرأها القائم على الاختبار شفهياً أو أن يقرأها الممتحَنون، واضحة وكاملة.

تتنوع المؤهلات المطلوبة من المسؤولين على الاختبارات وتكون المتطلبات في حالات الاختبارات الفردية واختبارات الذكاء والشخصية أكثر صرامة من متطلبات الاختبارات الجماعية. ويتعين على دور نشر وتوزيع الاختبارات التأكد قدر الإمكان من توفر الحد الأدنى من مؤهلات إجراء الاختبارات لدى الأفراد الراغبين في شراء الاختبارات منهم. ويُطلب من القائمين على الاختبارات أن يكونوا مسؤولين ليس عن الذي يجري أثناء فترة الاختبار فحسب بل عن التحضير للاختبار من حيث الناحية الأمنية والظروف المواتية قبل بدء الاختبار وعن التعامل الصحيح مع نتائج الاختبار بعد انتهائه. قد يحدث الغش في أي مرحلة من مراحل الاختبار (أي قبله أو أثناءه أو بعده) ولا يجب الخلط بين الغش

وبين البراعة في الاختبارات التي تعني أن بعض الأفراد موهوبون من حيث قدرتهم على تحقيق مستويات أداء عالية في جميع الاختبارات.

يجب الالتزام بطريقة إجراء الاختبار المحددة وخصوصاً في حالة الاختبارات القياسية. إلا أنه هنالك حالات تتطلب عدم التقيد بهذه الإجراءات مثل حالات اختبار الأفراد المعاقين الذين لا يمكنهم الالتزام بتعليمات الاختبار العادية. ولكن من المهم هنا مراعاة عدم مخالفة تعليمات الاختبارات إلا إذا كان الهدف الحصول على معلومات أكثر دقة فيما يخص بعض الممتحَنين، كما يجب مراعاة عدم مقارنة نتائج الحالات الإستثنائية مع المعايير المستخلصة من عينات المعايرة الطبيعية.

ويتم إجراء معظم الاختبارات الفردية بالطريقة التقليدية التي تتطلب ممتحن وواحد وشخص قائم على الاختبار، بينما تستخدم الاختبارات الجماعية الورقة وقلم الرصاص. وقد ازدادت مؤخراً شعبية الاختبارات التي تستعين بالحاسب وخاصة الاختبارات التكيفية (CAT) مثل اختبار التقييم المدرسي (SAT) واختبار سجل الخريجين (GRE) وسلسلة اختبارات القوات المسلحة للقابلية المهنية والتي تُجرى للآلاف من الأفراد داخل وخارج الولايات المتحدة.

وتعد الطريقة الأكثر استخداماً في وضع درجات الاختبارات الجماعية الموضوعية كالاختبارات التحصيلية واختبارات الاستعداد طريقة إعطاء درجة واحدة للإجابة الصحيحة ودرجة الصفر للإجابة الخاطئة. أما في اختبارات إعادة الترتيب والاختبارات المقالية فتعطى درجات مختلفة للإجابات بناءاً على دقتها أو شموليتها. وعلى الرغم من أن التصحيح من أثر التخمين لا يزال إجراء يرتبط بالعديد من الاختبارات، وكقاعدة عامة لا يعتبر استخدام مثل معادلات تصحيح التخمين هذه مع الاختبارات المدرسية خياراً صائباً.

يتم التعبيرعن درجات الاختبارات ذات المرجعية المحكية على شكل نسب مئوية بالنسبة إلى محك القبول أو مستوى الإتقان. أما في الاختبارات ذات المرجعية المعيارية فيتم تحويل الدرجات إلى معايير للعمر أو الصف أو معايير مئينية أو درجات معيارية تم الحصول عليها من نتائج عينة من الأفراد تمثل المجموعة البشرية المراد اختبارها. ويتوقع أن تكون عملية التعبير عن معدلات درجات الاختبارات عملية موضوعية إلا أنها في الواقع عملية ذاتية ترتبط بخبرة وتوقعات والتوجهات النظرية للشخص الذي يقوم بهذه العملية. وغالباً ما يتم تحويل درجات الاختبار في المدارس إلى درجات واستخدامها في تقرير نجاح أو رسوب الطالب.

يتم تدوين نتائج الاختبارات المدرسية في سجلات تراكمية أو في التقارير المدرسية. ويحدث شيء مشابه في مؤسسات التعليم العالي وبعض المؤسسات المهنية. أما في حالة الاختبارات النفسية والتشخيصية فيجب إعداد تقرير تفصيلي يضم معلومات عن شخصية الفرد الذي تم اختباره والسبب الذي أدى للاختبار ونتائج المقابلات والملاحظات والنتائج المحددة للاختبار النفسي بحد ذاته. ويختتم التقرير عادةً بقائمة من الاستنتاجات والتوصيات وذلك لغرض التقدير التالية أو طرق المعالجة أو غيرها من الإجراءات.

مصطلحات للمراجعة

راجع معاني المصطلحات المدرجة أدناه والتي تم استخدامها في هذا الفصل. الرجاء مراجعة فهرس التعريفات أو المعجم في حال وجود أي التباس حول معاني المصطلحات.

الاختبارات التحصيلية	السؤال المقالي
وضع الدرجات التحليلي	الاختبار ذو المستويات المرنة
اختبارات الاستعداد	وضع الدرجات الشمولي
تقنية حذف الكلمات من النصوص (cloze)	أسئلة المطابقة
الاختبار التكيفي المحوسب (CAT)	الاختبار متعدد المستويات
الوزن المرتبط بالثقة	أسئلة الخيارات المتعددة
وضع الدرجات عن طريق الاستبعاد	الاختبار الخارج عن المستوى
أسئلة إعادة الترتيب	التواصل
أسئلة الإجابات القصيرة	تصنيف الأهداف التعليمية
أدوات الدلالة المحددة	البراعة في الاختبارات
الكتيب الامتحاني اللولبي	أسئلة الصواب والخطأ

المراجع المقترحة

Aiken, L. R. (1991). Detecting, understanding, and controlling for cheating on tests. *Research in Higher Education, 32,* 725–736.

Airasian, P. W., & Terrasi, S. (1994). Test administration. In T. Husén & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (2nd ed.) (Vol. 11, pp. 6311–6315). Tarrytown, NY: Elsevier.

Golden, C. J., Sawicki, R. F., & Franzen, M. D. (1990). Test construction. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (2nd ed., pp. 21–40). New York: Pergamon Press.

Green, D. R. (1992). Test construction. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed.) (Vol. 4, pp. 1400–1406). New York: Macmillan.

Gronlund, N. E. (1993). *How to make achievement tests and assessments* (5th ed.). Boston: Allyn & Bacon.

Haladyna, T. M. (1994). *Developing and validating multiple-choice items.* Hillsdale, NJ: Erlbaum.

Lehman, R. (1994). Essays, scoring of. In T. Husén & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (2nd ed.) (Vol. 4, pp. 2018–2024). Tarrytown, NY: Elsevier.

Millman, J., & Greene, J. (1989). The specification and development of tests of achievement and ability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 335–366). New York: American Council on Education/Macmillan.

Nitko, A. J. (1989). Designing tests that are integrated with instruction. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 447–474). New York: Macmillan.

Robertson, G. J. (1990). A practical model for test development. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence & achievement* (pp. 62–85). New York: Guilford Press.

٤

الاختبارات التحصيلية القياسية

يتميز المجتمع الأمريكي، حسب وجهة نظر ديفيد مكليلاند (١٩٦١)، بأنه مجتمع يعشق الإنجاز ويستند إلى المبادئ البروتستانية المسيحية والمبادئ الرأسمالية التي تحث على العمل الدؤوب لتحقيق النجاح المادي والسعادة. وغالباً ما يتعلم الأولاد أن يعجبوا بالإنجازات التي يحققها الآخرون في مجالات الرياضة والترفيه والأعمال والعلوم (في بعض الأحيان) وأن يحاولوا أن يترسموا خطى هؤلاء الناجحين وأن يحققوا نجاحات أكبر وإنجازات أعظم. ولا ترتبط الرغبة بتحقيق الإنجازات بالخلفية الحضارية أو العائلية للفرد فحسب بل بالقابليات والإمكانيات التي يملكها هذا الشخص. وعلى الرغم من محاولاتنا الحثيثة، لن يتمكن إلا عدد قليل جداً منا من أن يصبحوا أبطالاً رياضيين أو فنانين مشاهير أو علماء عظماء، ويلعب الحظ كما تلعب الفرص دوراً في نجاحاتنا هذه ولكن يبقى السبب الرئيسي لعدم تمكن جميع الأفراد من تحقيق إنجازات عظيمة هو حقيقة أن الوراثة تحدد، إلى حدٍ ما، إمكانياتنا ومواهبنا. وحتى هذه الإمكانيات والمواهب لا يمكن أن تكون متساوية أذ قد نبرع في ناحيةٍ

ما بينما نفشل في آخرى. ويبقى التحدي الكبير بالنسبة لنا هو أن نعرف نقاط القوة التي نمتلكها ونقاط الضعف كذلك، وأن نبرز نقاط القوة التي يمكن أن تقودنا إلى النجاح. ولا يُطلب منا أن نمتلك إمكانيات خارقة لكي نصل إلى النجاح إذ يمكننا أن نعوض عن نقص الموهبة بالعمل وبذل الجهد. وعلى الرغم من أن العمل الشاق لن يعوض عن الإمكانيات التي هي دون المستوى إلا أن مستوىً من القدرة يعادل الحدود الطبيعية أو أعلى من ذلك بقليل يكفي لأن يقود أي فرد إلى التحصيل على مستوى عال إلى حد ما.

الأسس

يعد قياس مستوى التحصيل (وهو مستوى المعرفة أو المهارة أو الإنجاز الذي يحققه الفرد في مجالٍ معين) من أسهل عمليات قياس المتغيرات النفسية وأقلها تعقيداً. فقد تم، ولآلاف السنين، تقييم الإنجازات البشرية عن طريق ملاحظة السلوك البشري والإنتاجية البشرية وعن طريق الاختبارات والتجارب المعدة لهذا الغرض. وعلى الرغم من أن معظم اختبارات التحصيل كانت وعلى مر العصور اختبارات غير رسمية إلا أن أنواعاً رسمية من الاختبارات التحصيلية كاختبار التوظيف في المؤسسات الحكومية الصينية والذي تم استخدامه في الصين منذ أكثر من ٤٠٠٠ عام ظهرت بين الحين والآخر.

تعتبرالاختبارات التحصيلية القياسية أحد منتجات القرن العشرين، وهذا صحيح إلى حد كبير إلا أن عدداً من التطورات ظهرت في القرن التاسع عشر كنشر تعليمات لطريقة وضع وتحليل درجات الاختبارات (رسنيك، ١٩٨٢)، ونشر مقياس جورج فيشر للخط عام ١٨٦٤، واعتماد امتحانات ريجنت في ولاية نيويورك عام ١٨٦٥، وأبحاث جوزيف رايس (١٨٩٥) في تأثير التكرار على قدرة طلاب المدارس على التهجئة. وقد وضع رايس، الذي يعتبر رائد البحث العلمي في مجال التعليم والاختبارات القياسية (روس وستانلي، ١٩٥٤)، اختباراً مؤلفاً

من ٥٠ سؤالاً عن التهجئة واستخدمه في دراسته التي أجراها على ٣٣٠٠٠ طفل. وقد خلصت دراسة رايس إلى أن الأطفال يتعلمون تهجئة عدد من الكلمات خلال ١٥ دقيقة من التكرار يساوي ما يتعلمونه خلال ٤٠ دقيقة. إلا أنه، ومع الأسف، قادت دراسة رايس إلى زيادة الوقت الذي يقضيه المدرسون في تكرار الكلمات مع الطلاب وليس في تقليصه. ووضع رايس اختبارات للقدرة اللغوية والحسابية كجزء من أبحاثه في تأثير طريقة التدريس على معرفة الطلاب في هذه المواد.

وقد جاء تطوير الاختبارات التحصيلية القياسية في الربع الأول من القرن العشرين ليكمل جهود الرواد في مجال الاختبارات كرايس وغيره. ففي عام ١٩٠٨ ظهر اختبار ك. ل. ستون في مجال الحساب وهو اختبار للعمليات الحسابية الأساسية والاستدلال Reasoning الحسابي (١٩٠٨)، وفي العام التالي ظهرت سلسلة الاختبارات الحسابية التي وضعها س. أ. كورتيس (١٩٠٩). أما في عام ١٩١٢ فقد اكتشف ستارش وأليوت أن الإتفاق بين المدرسين في الدرجات التي أعطوها للاختبارات المقالية كان مخيباً للآمال مما أدى إلى تدعيم أسس الاختبارات الموضوعية (راجع ستالناكر، ١٩٥١). كما ساهم تطوير اختبارات الذكاء الجماعية والفردية في فترة الحرب العالمية الأولى في تهيئة المناخ لازدهار ونمو الاختبارات القياسية. وتم في العشرينيات نشر أول اختبارين قياسيين وهما اختبار ستانفورد للتحصيل واختبار مدارس آيوا الثانوية للمحتوى. أما عام ١٩٢٦، فقد شهد استبدال مجلس اختبارات القبول في الكليات والجامعات للاختبارات المقالية التي كان يستخدمها مع الطلاب الراغبين بالالتحاق بمؤسسات التعليم العالي باختبار متعدد الخيارات هو اختبار الاستعداد المدرسي (SAT). وقد أسهم استخدام هذا النوع من الاختبارات، بالإضافة إلى تطوير آلات تساعد في وضع الدرجات أوتوماتيكياً، في تحفيز تقييم تحصيل الطلاب عن طريق الاختبارات القياسية. كما توسعت الاختبارات

القياسية، منذ الحرب العالمية الثانية، وذلك من اختبارات تتناول مادة واحدة إلى اختبارات أكثر شمولاً تغطي العلوم الإنسانية أو العلوم الطبيعية بأكملها.

ومن الخداع القول إن التطور الهائل الذي شهده مجال الاختبارات القياسية في الولايات المتحدة يُعزى إلى التطور الفكري والتقني الذي شهدته البلاد، إذ يوضح ليفاين (١٩٧٦) أن شعبية الاختبارات التحصيلية القياسية في الولايات المتحدة ارتفعت لأن جميع الأطراف التي دخلت في معركة سياسة حول قضية التعليم وجدت في نتائج الاختبارات القياسية دعماً لوجهات نظرها. ومنذ ذلك الوقت وحتى الآن استُخدمت نتائج الاختبارات القياسية في النقاشات الحادة على الساحة السياسية. وكجزءٍ من الجهود الدؤوبة لرفع المستوى التعليمي لمواطني الولايات المتحدة بكامل أطيافهم السياسية، يستخدم العاملون في مجال السياسة نتائج الاختبارات التحصيلية القياسية لدعم مشاريعهم التي تستدعي تدخلاً أقل أو أكثر من الحكومة في التعليم أو رصد ميزانية أكبر للتعليم أو تجربة برامج تعليم خاصة أو غيرها. وتعد الإجراءات التي طبقتها الدولة في السنوات الأخيرة، كربط تخرج طلاب الثانوية بنجاحهم في اجتياز اختبار في عدد من المهارات الأساسية وإجراء تقييمات على مستوى الولاية وتمويل عدد من المواد والمشاريع واعتبار المدرسين مسؤولين عن التحصيل العلمي للطلاب، إجراءات تدل على المشاركة السياسية الواسعة في مجال التعليم.

لا تزال معظم الاختبارات التحصيلية القياسية اختبارات موضوعية متعددة الخيارات، إلا أن الحملة الدعائية الهائلة ضد هذا النوع من الاختبارات والتي تحدثت عن تدني مستوى الكتابة ومهارات التفكيرعند الطلاب وعدم قدرتهم على نقل المهارات والمعارف التي حصلوا عليها داخل الفصول الدراسية إلى العالم الحقيقي، بالإضافة إلى بُعد المادة الدرسية عن الواقع، أدت إلى أعادة النظر في شكل ومحتوى الاختبارات التعليمية. وتتضمن التغييرات المقترحة في هذا

السياق: وضع اختبارات تقيس الأهداف التعليمية العليا كالإستيعاب والفهم، واختبارات تقيس مهارات الدراسة والعمل، واختبارات صممت خصيصاً لمعلومات كتاب أو برنامج تعليمي معين، واختبارات مرجعية المحك، واختبارات تقيس إمكانيات التعبير والأداء الأخرى، واختبارات لقواعد التجارب العملية، وملفات عمل طلابية تتضمن عينات من عمل الطلاب في مختلف المواد (لين، ١٩٩٢). وتبقى "صناعة الاختبارات"، على الرغم من الانتقادات التي تتعرض لها عملية الاختبار، صناعةً لا تزال في نمو مستمر من حيث تنوع نشاطاتها ومن حيث عدد الاختبارات التي تباع يوماً بعد يوم. وبحلول عام ١٩٦٦، اتسع استخدام الاختبارات التحصيلية القياسية بشكل وصلت فيه مبيعات الاختبارات لسنين الحضانة إلى الثانوية العامة (K-12) إلى مبلغ قدر ب ١٧٨,٧ مليون دولار (اتحاد الناشرين الأمريكيين، تواصل شخصي).

دلالات الاختبارات التحصيلية:

إن الغرض الأساسي لإجراء الاختبارات التحصيلية هو تقييم مستوى فردٍ ما من حيث تمكنه أو معرفته أو مهارته أو فهمه ضمن مجالٍ إدراكي ما، وغالباً ما تعد هذه الاختبارات إجراءات تقييم خلاصية تستخدم في نهاية الوحدة الدراسية أو المقرر أو السنة الدراسية لتحدد درجة ضلوع الطالب في الأهداف التعليمية التي تم تغطيتها. ويُعطى الممتحن، بناءاً على درجاته في اختبار التحصيل، درجة نهائية أو ترفيعاً وظيفياً أو شهادة تقر بحصوله على درجة من الكفاءة في المادة التي تم اختباره فيها. كما يمكن أن تستخدم درجات اختبارات التحصيل لتوزيع الطلاب على مستويات مقرر معين أو منحهم مكاناً متقدماً في مادة ما أو منحهم ساعات معتمدة في مقرر ما دون حضور المقرر. وبالإضافة إلى استخدامها كإجراءات تقييم تراكمية Summative evaluation، يمكن استخدام اختبارات التحصيل كإجراءات تقييم تكوينية Formative evaluation

تساعد في إعطاء معلومات تساعد على تعديل طريقة تدريس فرد ما خلال مقرر لا يزال مستمراً. كما تساهم درجات اختبارات التحصيل في تشخيص نقاط الضعف والقوة لدى الأفراد المختلفين في مادة ما أو في تحديد مجالات العجز التعليمية وحالات الاضطرابات السريرية.

وغالباً ما تكون بيانات اختبارات التحصيل لفرد ما متاحة للمدرسين وإداريي المدرسة والطلاب أنفسهم وأولياء أمورهم لتسهيل التعاون بين المدرسة والبيت بما فيه مصلحة الطلاب. كما يمكن لاختبارات التحصيل أن تحفز الطلاب على العمل على رفع مستواهم في مادة ما أو أن تزيد من رغبتهم في تعلم المزيد.

ولا تقتصر استخدامات الاختبارات التحصيلية في مجال التعليم على المساعدة في إتخاذ قرارات بشأن الأفراد بل تساعد في تقييم المناهج والبرامج التعليمية وحتى المدرسين أنفسهم حيث توفر درجات الاختبارات التحصيلية للعاملين في مجال التدريس والإدارة التعليمية معلومات قيمة تساعدهم في التخطيط للمناهج وتقييم فعالية أدائهم وإجراء تعديلات عند اللزوم. كما تساهم المعلومات التراكمية التي يتم جمعها من درجات الاختبارات التحصيلية في إتخاذ قرارات متعلقة باختيار وتوزيع وترقية ونقل وعدم فصل العديد من العاملين في مجال التدريس والإدراة التعليمية، كما يمكن أن تساعد في رصد وتوزيع الموارد المالية والتعليمية وتجربة برامج تعليمية جديدة أو الاستمرار في البرامج التعليمية القائمة أو الاستغناء عن بعضها وغير ذلك من الأمور المتعلقة بالسياسة التعليمية.

لا يقتصر استخدام الاختبارات التحصيلية على مجال المدرسة أو التعليم فقط بل يتعداه ليدخل، وبشكل واسع، في مجالات الصناعة والتوظيف والدولة والقوات المسلحة وغيرها من المؤسسات التي تعمل في مجالات الإنتاج أو الخدمات. وتساعد اختبارات التحصيل في هذا السياق في تقييم المتقدمين

بطلبات توظيف أو الموظفين بغرض إتخاذ قرارات بشأن تعيينهم أو ترقيتهم أو توجيههم. ويطلق على هذا النوع من الاختبارات اسم الاختبارات المهنية والتي تساعد مع اختبارات الاستعداد والشخصية والأداء والمعلومات الشخصية عن الأفراد في توفير المعلومات اللازمة لإتخاذ قرارات فردية أو جماعية في مجال التوظيف.

الاختبارات القياسية أم الاختبارات التي يضعها المدرسون؟

على الرغم من أن مبيعات الاختبارات التحصيلية القياسية تزيد عن مبيعات أي نوع من الاختبارات النفسية أو التعليمية، يبقى عدد الاختبارات التحصيلية التي يضعها المدرسون أكبر بكثير من عدد اختبارات التحصيل القياسية التي يضعها الأخصائيون ومحترفو وضع الاختبارات. ويختلف محتوى الاختبارات التحصيلية القياسية عن تلك التي يضعها مدرسو المقرارات. فعلى حين تركز الاختبارات التي يضعها المدرسون على الأهداف التعليمية للمدرس أو السنة الدراسية، يكون التركيز في الاختبارات القياسية على الأهداف التعليمية لعدد من المدارس أو لمنطقة تعليمية كاملة. وعلى عكس المدرسين الذين لم يتلقوا تدريباً خاصاً في مجال وضع وتصمميم الاختبارات والذين يضعون اختبارات تحصيل خاصة بوضع ما وذات مدة صلاحية محدودة، ينتمي واضعو الاختبارات القياسية إلى فئة من الخبراء في المواد العلمية المختلفة تم تدريبهم على وضع الاختبارات أو تضافرت جهودهم مع خبراء التقييم لوضع اختبارات قياسية لا حدود لمدة صلاحيتها. وبينما تركز اختبارات المدرسين على معلوماتٍ شديدة التحديد، تُعنى اختبارات التحصيل القياسية بقياس القدرة على التفكير والإستيعاب. وأخيراً تتم مقارنة درجات الاختبارات التحصيلية التي يضعها المدرسون مع توقعات المدرس الشخصية ومقاييسه الخاصة وأهدافه، بالإضافة إلى ما يتذكره عن أداء الطلاب في السنوات السابقة. أما في حالة الاختبارات القياسية ذات المرجعية المعيارية فغالباً ما تتوفر معايير تُرفق بالاختبار تساعد

في فهم درجات الاختبارات وتحديد مصير الطالب من حيث الرسوب أو النجاح. أما في حالة الاختبارات المقننة ذات المرجعية المحكية فلا تتم مقارنة النتائج بمعايير أو معدل درجات عينة القياس بل بالرجوع إلى مجموعة من مقاييس أو محكات أداء تم الاتفاق عليها.

وتعد إمكانية مقارنة أداء فردٍ ما مع أفراد غيره ينتمون إلى مستوى القدرة ذاته ميزة أخرى لاختبارات التحصيل القياسية التي تمت معايرتها مع اختبارات ذكاء أو اختبارات عامة للاستعداد الأكاديمي. فعلى سبيل المثال يمكن مقارنة درجات الطبعة السابعة من اختبار متروبوليتان للتحصيل مع درجات الطبعة السادسة لاختبار أوتيس – لينون للقدرة المدرسية، أو مقارنة درجات اختبارات أيوا للمهارات الأساسية أو اختبارات التحصيل والكفاءة أو اختبارات أيوا للإرتقاء العلمي مع درجات اختبارات القدرات الإدراكية.

ومن المهم أن نلاحظ أن اعتبار الاختبارات ذات المرجعية المعيارية والاختبارات ذات المرجعية المحكية اختبارات تعمل بشكلين مختلفين هو اعتبار مضلِّل. إذ أن أي اختبار للتحصيل يمكن أن يعطينا معلومات تفصيلية عن أداء شخصٍ ما ليس بالمقارنة مع مجموعة من المعايير والأهداف التعليمية الثابتة التي تم اعتمادها مسبقاً بل بالمقارنة مع غيره من الأفراد الذين ينتمون للمرحلة العمرية أو الدراسية ذاتها.

تمتاز الاختبارات التحصيلية القياسية والاختبارات التحصيلية التي يضعها المدرسون بالتكامل وليس بالتعارض وذلك من حيث التطبيقات والمزايا على الرغم من الإختلافات الظاهرة بينها، فلكلا النوعين استخداماتها المهمة بشكل عام (في حالة الاختبارات القياسية) وبشكل خاص (في حالة اختبارات المدرسين). ويعطي كلا النوعان معلومات عن تحصيل الطلاب بالنسبة إلى الأهداف التعليمية للمدرسة أو السنة الدراسية أو لمجموعة من المدارس. وتعد الاختبارات القياسية

أكثر ثباتاً من الاختبارات التي يضعها المدرسون ولذلك يعتمد عادةً على هذا النوع من الاختبارات عند المقارنة بين التحصيل العلمي للطلاب أو عند توزيعهم على المقررات والمناهج المختلفة أو عند تقييم مناهج المدارس المختلفة. ولا يقتصر عمل اختبارات التحصيل القياسية وتلك التي يضعها المدرسون على تقييم مستوى تحصيل الطلاب، إذ توفر هذه الاختبارات معلومات تشخيصية تساعد على إعداد برامج تعليمية خاصة لكل طالب وللطلاب ذوي الإحتياجات الخاصة كما تساعد على توزيع الطلاب على مستويات تعليمية مختلفة. ومن الجدير بالذكر أن الاختبارات القياسية تفيد أكثر في العمليات المذكورة أعلاه.

مراجع للمعلومات خاصة بالاختبارات التحصيلية القياسية:

يضم الفصل الأول من هذا الكتاب معلومات عن المراجع المتوفرة عن الاختبارات التحصيلية بشكل خاص والاختبارات القياسية ولعل أهم هذه المراجع كتاب الاختبارات المنشورة ٤ Tests in Print IV (مورفي وكونولي وإيمبارا، ١٩٩٤) وكتاب اختبارات Tests (سويتلاند وكيسر، ١٩٩١). ويمكن الحصول على معلومات عن الاختبارات المنشورة من دليل الاختبارات المنشورة (راجع الملحق ب). كما يضم برنامج F-8 الذي يحتويه القرص المرفق بهذا الكتاب قائمة بأسماء ومستويات وناشري ومصادر النقد لمئاتٍ من الاختبارات التحصيلية المتوفرة في الأسواق.

ومن المفيد قبل اختيار اختبار تحصيلي ما الحصول على عينة من هذا الاختبار ودليل المستخدم الخاص به. كما يفيد الإطلاع على رأي النقاد في الاختبار والذي يمكن الحصول عليه بالعودة إلى كتابي: الكتاب السنوي للمقاييس العقلية The Mental Measurements yearbooks (كونولي وإيمبارا، ١٩٩٥ والطبعات السابقة) وكتاب مقالات نقدية عن الاختبارات Test Critiques (كيسر وسويتلاند، ١٩٨٤-١٩٩٤) ومن ثم إتخاذ قرار بشأن الاختبار من حيث مناسبته للأغراض المراد تحقيقها وللمجموعة المراد اختبارها وللظروف التي سيجرى

الاختبار ضمنها. ويُعنى النقاد غالباً بوصف محتوى الاختبار وطريقة إجرائه ووضع درجاته وطريقة قياسه ووضع معاييره بالإضافة إلى إعطاء معلومات عن ثباته وصدقه. ويحتوي الملحق ٤-١ على نموذج عن مقال نقدي عن أحد اختبارات التحصيل التسلسية وهو اختبار مستويات التحصيل للمنهج التدرجي.

وسنتطرق في ما بقي من هذا الفصل إلى الحديث عن بطاريات التحصيل ثم ننتقل إلى اختبارات المسح والتشخيص والتنبؤ للمواد المدرسية ونختم بالحديث عن الاختبارات التحصيلية في مجال الحرف والصناعات اليدوية.

بطاريات الاختبارات التحصيلية

تقيس بطاريات الاختبارات التحصيلية المهارات الرئيسة عند الطلاب كالقراءة والقواعد والحساب (الرياضيات)، وفي حالة طلاب المراحل المتوسطة والثانوية تقيس كذلك مهارات الدراسة والعلوم الطبيعية والإنسانية. وتعد بطارية اختبار ستانفورد للتحصيل واختبارات أيوا للمهارات الأساسية (ITBS) من أول البطاريات التحصيلية التي تم نشرها وأكثرها شعبيةً واستخداماً. ويضم برنامج F-8 الذي يحتويه القرص المرفق بهذا الكتاب قائمة بأسماء العشرات من الاختبارات التحصيلية المستخدمة في الولايات المتحدة وغيرها من دول العالم. ومن أكثر هذه الاختبارات شعبية في الولايات المتحدة الاختبارات التالية:

- اختبارات كاليفورنيا للتحصيل (CAT)
- الاختبار الشامل للمهارات الأساسية (CTBS)
- اختبار أيوا للتطور العلمي (ITED)
- اختبار متروبوليتان التحصيلي (MAT)
- سلسلة SRA التحصيلية
- اختبارات التحصيل والكفاءة (TAP)

تُستخدم جميع اختبارات التحصيل المتعددة المستويات المذكورة أعلاه في اختبار الطلاب من عمر الحضانة إلى الثانوية العامة باستثناء اختبار ITBS الذي يتوقف عند الصف الثامن واختباري ITED وTAP المعدين للطلاب بين الصف التاسع والثالث الثانوي. كما يتوفر عدد من الاختبارات التي تستخدم مع أفراد من شتى مراحل العمر كاختبار التحصيل الواسع المدى ٣ (للأشخاص بين ٥ و٧٥ عاماً) وبطارية اختبارات وودكوك -جونسون النفسية- التعليمية (النسخة المعدلة) والذي يستخدم مع الأفراد الذين تتراوح أعمارهم بين ٢ و٩٠ عاماً، واختبار كوفمان للمهارات الأكاديمية الوظيفية (للأفراد بين ١٥ و٨٥ عاما وما فوقً).

وهناك برنامج اختبار نموذجي يتم فيه اختبار الطلاب في المدارس الإبتدائية في الخريف والربيع باستخدام أحد بطاريات التحصيل وذلك لتقييم تقدمهم الدراسي. ونظراً لأن قياس الاختبارات المختلفة التي تتألف منها البطارية تم بالنسبة إلى العينة البشرية ذاتها؛ فإنه يمكن مقارنة نتائج الفرد في اختبارٍ ما مع نتائج غيره من الأفراد أو مقارنة نتائج الفرد في اختبار ما مع نتائجه في الاختبارات الأخرى. وعلى الرغم من أنه من الممكن مقارنة الدرجات ووضع رسم بياني يوضح تقدم مستوى الفرد عبر السنوات، يجب الحرص عند افتراض أن مجموعات الدرجات المختلفة التي تم عليها قياس الاختبار كانت متكافئة في القدرات، وإذا كان هذا الافتراض غير صحيح فلا يمكن مقارنة درجات الفرد النهائية (مكافئات العمر والصف والرتب المئينية والدرجات القياسية) لعمر معين أو سنة دراسية معينة بشكلٍ مباشر. ومن المهم كذلك التأكيد على أن العلاقة بين العمر والوسط الحسابي والوسيط لدرجات الاختبارات التحصيلية هي علاقة منحنية وليست خطية - طولية يصغر حجم وحدات الدرجات فيها مع تقدم الطالب في السنوات الدراسية. كما يحقق الطلاب ذوي مستويات القدرة العالية قفزات في التحصيل لا يحققها الطلاب ذوي مستويات القدرة المتدنية.

وتستخدم بطاريات اختبارات التحصيل في المرحلة الإبتدائية أكثر من المراحل المتوسطة والثانوية بسبب التطور الكبير الذي يطرأ على الطلاب في هذه المرحلة وبسبب التقارب الشديد بين مناهج المدارس الإبتدائية. وكما ذكرنا سابقاً، وكما يوضح الشكل ٤-١، تعنى بطاريات اختبارات التحصيل بتقييم المهارات الأساسية لدى الطلاب كالقراءة (مفردات واستيعاب) والمهارات اللغوية (تقنيات الكتابة والتعبير) والرياضيات (الحساب والمفاهيم والتطبيقات) ويضاف عليها في مراحل الدراسة المتوسطة والثانوية فقرات في التهجئة والمهارات الدراسية والعلوم الطبيعية الدراسات الاجتماعية.

وتتألف معظم بطاريات اختبارات التحصيل من أسئلة اختبارات متعددة إلا أن بعض هذه الاختبارات (كالطبعة التاسعة من سلسلة اختبارات ستانفورد للتحصيل) تتضمن بعض الأسئلة التي تطلب من الطلاب كتابة مسودة مقال عن موضوع معين أو الإجابة على أسئلة تتطلب إجابات مفتوحة (وذلك في الرياضيات والعلوم والعلوم الإنسانية).

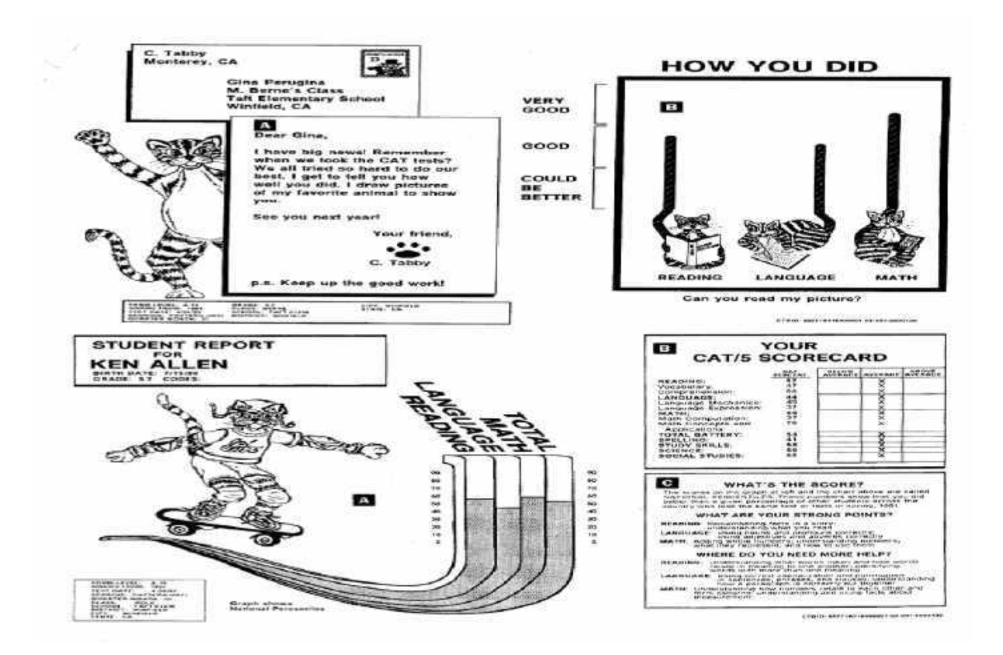
ولا يقتصر استخدام بطاريات اختبارات التحصيل على طلاب المدارس فقط بل، وكما ذكرنا سابقاً، يوجد عدد من هذه الاختبارات التي تصلح للاستخدام مع الأفراد من مختلف الأعمار مثل اختبار التحصيل الواسع المدى ٣، وبطارية اختبارات وودكوك - جونسون النفسية Woodcock - Johnson Rhsychoed التعليمية (النسخة المعدلة)، واختبار كوفمان للمهارات الأكاديمية الوظيفية. وتضاف إلى هذه القائمة ثلاثة بطاريات اختبارات صممت لتقدير الحالة التعليمية لليافعين الذين لم يتخرجوا من المدارس الثانوية، وهذه الاختبارات هي: امتحان التعليم الأساسي للبالغين (ABLE) واختبارات التعليم الأساسي للبالغين (TABE) واختبارات التطور العلمي العامة (GED). ومثل اختبارات التعليم الأساسية فإنه على حين يركز ABLE وTABE على المهارات التعليمية العامة

كالقراءة واللغة والرياضيات، يتضمن اختبار GED (والذي يستغرق ٧ ساعات كاملة) أسئلة الخيارات المتعددة التي تهدف إلى تقييم فهم الممتَحنين للمفاهيم المعرفية العامة التي تتم تغطيتها في مرحلة الدراسة الثانوية. وبسبب هذا التقارب بين محتويات اختبار GED ومناهج المدارس الثانوية تمنح عدد من الكليات ومؤسسات الأعمال والخدمات العسكرية شهادات تخرج للطلاب الذين ينجحون في تحقيق معدل درجات معين في اختبارات GED.

وأخيراً على الرغم من أن تسويق اختبارات القبول الجامعي في الولايات المتحدة (ACT) يتم على أساس كون هذه الاختبارات معدة للقبول وليست بطاريات اختبارات تحصيلية، تقيس هذه الاختبارات التحصيل في الرياضيات واللغة الإنجليزية والقراءة والتفكير العلمي. يشبه اختبار ACT اختبارات الاستعداد (التي سنتحدث عنها بالتفصيل في الفصل السادس من الكتاب) من حيث تركيزه على قاعدة واسعة من المعلومات لا ترتبط بمعلومات الدراسة الثانوية ارتباطاً وثيقاً كما هو حال معظم بطاريات التحصيل. ويتم إجراء اختبار القبول في جامعات وكليات الولايات المتحدة (ACT) أربع مرات في السنة وفي أكثر من ٤٠٠٠ مركز للامتحانات داخل الولايات المتحدة وخارجها بهدف تقييم طلاب المدارس في النواحي الأربعة المذكورة أعلاه. ثم يتم جمع نتائج الطالب في المهارات الأربعة ويتراوح المجموع بين ١ و٣٦ بمعدل وسطي هو ١٨، بينما تتراوح درجات الاختبارات الفرعية لكل مهارة بين ١ و١٨ بمعدل وسطي هو ٩، ويعد هذا الاختبار، اختباراً جيداً لا يسبقه في الشعبية، كاختبار للقبول الجامعي، إلا اختبار SAT.

الشكل ٤-١

اختبار كاليفورنيا التحصيلي، الطبعة الخامسة (CAT/5)، تقرير مدرسي لطلاب المراحل الإبتدائية والمتوسطة. (حقوق الطبع محفوظة ١٩٩٢، تمت إعادة الطبع بموافقة شركات ماكغرو - هيل المتحدة)



اختبارات المادة الواحدة

يساعد إجراء بطاريات التحصيل التي تعطي فكرة عامة عن مهارات المتحَن التعليمية والمعرفية التي تحتل الأولوية نموذجياً في برنامج الاختبار المدرسي. إن مثل بطارية الاختبار هذه والخاصة بالمهارات الأساسية تتبع باختبارات المادة الواحدة التي تغطي مجالاً واحداً أو مجالات عدة محددة وذلك لإعطاء صورة أفضل عن إمكانيات الشخص في مواضيع محددة. وعلى عكس بطاريات التحصيل التي غالباً ما تغطي مواد عدة مما يقلل من التركيز على المواد المختلفة، يتألف اختبار المادة الواحدة من فقرات اختبارية دقيقة وعميقة

وعديدة. ولهذا السبب تعد هذه الاختبارات أكثر ثباتاً وصدقاً من بطاريات التحصيل كما أنها أكثر تعبيراً عن قدرات الفرد في مجالٍ تعليمي ما. إلا أن المشكلة الرئيسة هي أنه، وكما هو الحال في البطاريات التحصيلية، يتعذر مقارنة درجات فردٍ ما مع فردٍ آخر بسبب اعتماد عملية القياس على عينات مختلفة من الأفراد.

يتوفر تجارياً اختبارات مادة واحدة في المجالات التالية: القراءة واللغة الإنجليزية والعلوم الإنسانية واللغات الأجنبية والرياضيات والعلوم وغيرها من المواد المدرسية. وتعد معظم هذه الاختبارات، اختباراتٍ ذات مرجعية معيارية حتى أن بعض الاختبارات ذات المرجعية المحكيةكاختبارات كاليفورنيا التشخيصية والتي، على الرغم من كونها اختبارات موضوعية، تستخدم المعايير شأنها كشأن الاختبارات ذات المرجعية المعيارية. كما تتوفر أعداد هائلة أخرى من الاختبارات التجارية أو التي يسوقها الخبراء في مجال الاختبار والتي تساعد في اختيار وتوزيع الطلاب والموظفين في المؤسسات التعليمية والرسمية. وسوف يتحدث الباب التالي من هذا الفصل عن أنواع هذه الاختبارات واستخداماتها. ومن الجدير بالذكر أن برنامج F-8 المرافق للكتاب يضم قائمة مفصلة بأسماء اختبارات المادة الواحدة المتوفرة في الأسواق والتي يمكن الحصول على معلومات أدق عنها من الكتب التالية: الاختبارات المطبوعة ٤ (مورفي، كونولي وإيمبارا، ١٩٩٤)، والكتب السنوية للمقاييس العقلية، وكتاب اختبارات (سويتلاند وكيسير، ١٩٩١) وكتاب مقالات نقدية عن الاختبارات.

اختبارات القراءة

تعزى معظم المشاكل التي يعاني منها طلاب المدارس إلى صعوبات تواجههم في قراءة النصوص وعلى الرغم من أنه لا توجد تقسيمات واضحة لأنواع اختبارات القراءة يمكن تقسيمها إلى اختبارات مسح (أو تقييم عام للقدرة)

واختبارات تشخيص واختبارات جاهزية. وتهدف اختبارات المسح إلى تقييم إمكانيات القراءة العامة لدى الفرد وتتضمن غالباً عدداً من الأجزاء التي تحتوي على عدد من المفردات والمقاطع النصية التي يطلب من الطالب الإجابة عن أسئلة تتعلق بها. وتعد بعض اختبارات المسح مناسبة لطلاب مرحلة ما قبل المدرسة بينما تناسب اختبارات أخرى طلاب المرحلة الإبتدائية أو حتى طلاب المراحل الوسطى والثانوية. إلا أن عدداً من هذه الاختبارات (كاختبارات غايتس – ماكجينايتي للقراءة والتي تعد من أكثر اختبارات المسح شعبية) تتألف من عدة مستويات ويمكن استخدامها مع الطلاب من سن الحضانة إلى الثانوية العامة. كما تتوفر اختبارات أكثر محدودية من حيث أعمار الطلاب التي يمكن استخدامها معهم كاختبار نلسون – دني للقراءة والذي يستخدم مع طلاب الصف التاسع وحتى الصف السادس عشر[١]. ويحتوي برنامج F-6 على نموذج من الاختبارات التي تقيس سرعة القراءة والإستيعاب.

يستخدم مصطلح "اختبار تشخيصي" كثيراً في مجال الاختبارات وغالباً على نحو غير دقيق. تقيم اختبارات القراءة التشخيصية المعدة بعناية القدرات المعقدة والأساسية التي يحتاجها الفرد ليصبح ضليعاً في مجال القراءة. فعلى سبيل المثال تتضمن اختبارات القراءة التشخيصية المتعددة المستويات MAT6 عدداً من الاختبارات الفرعية التي تقيس:

- التمييز السمعي والبصري.
- وقدرة التعرف على الحروف.
- وقدرة التعرف على المفردات الأساسية.

---

(١) يعتبر بعض خبراء التعليم في الولايات المتحدة سنوات الدراسة الجامعية الأربعة استمراراً للتعليم المدرسي مما يعني أن سنة التخرج هي سنة الدراسة السادسة عشرة. (ملاحظة المترجم).

- والتعرف على الوحدات الصوتية والكتابية: الحروف الساكنة وحروف العلة.
- واستخراج معاني المفردات من السياق.
- واستخدام أجزاء الكلمات (اللواحق والسوابق) لاستنتاج معنى الكلمة.
- ودرجة الإستيعاب.
- والقراءة للمعنى العام والقراءة لاستخراج معلومات محددة.
- وفهم واستيعاب المادة النصية.

وتُسهم درجات الطالب في الاختبارات الفرعية المذكورة أعلاه في تقييم نقاط ضعف وقوة الطلاب (من الحضانة إلى الصف التاسع) في القراءة.

وتساهم اختبارات الجاهزية للقراءة (والتي تشبه اختبارات التشخيص إلى حد كبير إلا أنها تركز على التنبؤ بقدرة الفرد على القراءة) في تحديد ما إذا كان الطالب قد وصل إلى درجة كافية من النضج تؤهله لبدء تعلم القراءة. وتعد القراءة مهارة إدراكية وإدراكية - حسية معقدة تتطلب نضجاً جسدياً ونفسياً معيناً من الطالب لا يصل إليه الأطفال في العمر نفسه. وتصل البنات إلى هذا العمر قبل الصبيان ويكون ذلك غالباً بين عمري السادسة والسادسة وستة أشهر. وغالباً ما يتمكن الأطفال الأذكياء والذين لا يعانون من الديسلكسيا أو غيرها من الاضطرابات التي تؤثر على القراءة من إتقان مهارة القراءة وهم أصغر من غيرهم بكثير، ولهذا السبب تساهم اختبارات الذكاء العام كاختبار ستانفورد - بينيه ومقياس ويكسلر لذكاء الأطفال في تحديد جاهزية الأطفال للقراءة شأنها كشأن الاختبارات المعدة خصيصاً لهذا الغرض.

وبالرغم من أهمية القراءة كمهارة أكاديمية محورية، يوجد عدد من المهارات الأخرى كالكتابة والكلام والحساب التي لا تقل أهميتها عن مهارة القراءة. ولذلك تتوفر اختبارات مثل اختبارات متروبوليتان للجاهزية واختبار الجاهزية المدرسية

والتي تركز على المهارات العددية والتفكير والخط والتهجئة بالإضافة إلى المهارات المربطة بالقراءة (كالتمييز السمعي والبصري والتعرف على الحروف والربط بين الحروف والأصوات وغيرها).

اختبارات اللغة

بدأت معظم الجامعات في الولايات المتحدة ومنذ الستينيات بإلغاء أو تقليص أهمية واحدة من متطلبات القبول لديها وهي متطلب اللغة الأجنبية. وعلى الرغم من هذا، يدرس الطلاب في المدارس الثانوية عدداً من اللغات الأجنبية كالإسبانية أو الفرنسية أو الألمانية أو الإيطالية أو الصينية أو اليابانية أو العبرية أو اللاتينية أو الأندونيسية أو الهوسا Hausa. ويمكن للطالب الذي درس لغة أجنبية ما في المدرسة الثانوية لمدة عامين أن يُعفى من مقرر اللغة الأجنبية في الجامعة أو أن يعطى عدد الساعات المعتمدة لمقرر اللغة الأجنبية في حال إجتيازه بنجاح لواحد من الاختبارات التي تستخدمها الجامعة لتقييم مستوى الطلاب في اللغات المختلفة. ومن بين هذه الاختبارات: اختبار برنامج القبول في الصفوف اللغوية المتقدمة (والذي يتألف من اختبارات منفصلة للغة والأدب في الفرنسية والإسبانية والألمانية واللاتينية) واختبار CELP للمواد الدراسية واختبارالساعة الواحدة لمجلس الكليات لتقييم التحصيل (والذي يتألف من اختبار لمهارات الإستماع والقراءة في الفرنسية والألمانية والإسبانية) واختبار مجلس الكليات SAT II للغات. ويمكن اختبار طلاب السنة الأخيرة وأعضاء الهيئة التدريسية المختصين في الفرنسية والإسبانية والألمانية باستخدام اختبار NTE للنواحي العلمية التخصصية في تلك اللغة. تركز معظم اختبارات اللغة على تقييم مهارتي القراءة والاستماع إلا أن الاختبارات التي أعدها مركز علم اللغات التطبيقي للغات الصينية والعبرية والبرتغالية والإندونيسية والهوسا ولغات أخرى تقيس قدرة الطالب على تكلم اللغة كذلك. وتتوفر اختبارات (كاختبار الاستعداد في اللغات الأجنبية) يمكنها التنبؤ بقدرة الطالب الذي تعد اللغة الإنجليزية لغته الأم على تعلم لغة أجنبية ما بسهولة وسرعة.

لقد أسهمت صفة العالمية التي ارتبطت باللغة الإنجليزية مؤخراً في ابتعاد الناطقين باللغة الإنجليزية (كالأمريكيين والبريطانيين) عن تعلم اللغات الأجنبية. وكغيرها من اللغات الأخرى تعد اللغة الإنجليزية وسيلة من وسائل التواصل التي تستخدم عدداً من المفاهيم التي يجب أن يفهمها الفرد ويستخدمها في التعبير عن نفسه إذا أراد أن يكون عضواً فعالاً في المجتمع. وبناءاً على هذا، تساعد اختبارات كاختبار بوهيم Boehm للمفاهيم الأساسية في تحديد المشاكل التي يمكن أن يعاني منها الفرد أثناء تعلمه لهذه المفاهيم (راجع الشكل ٤-٢).

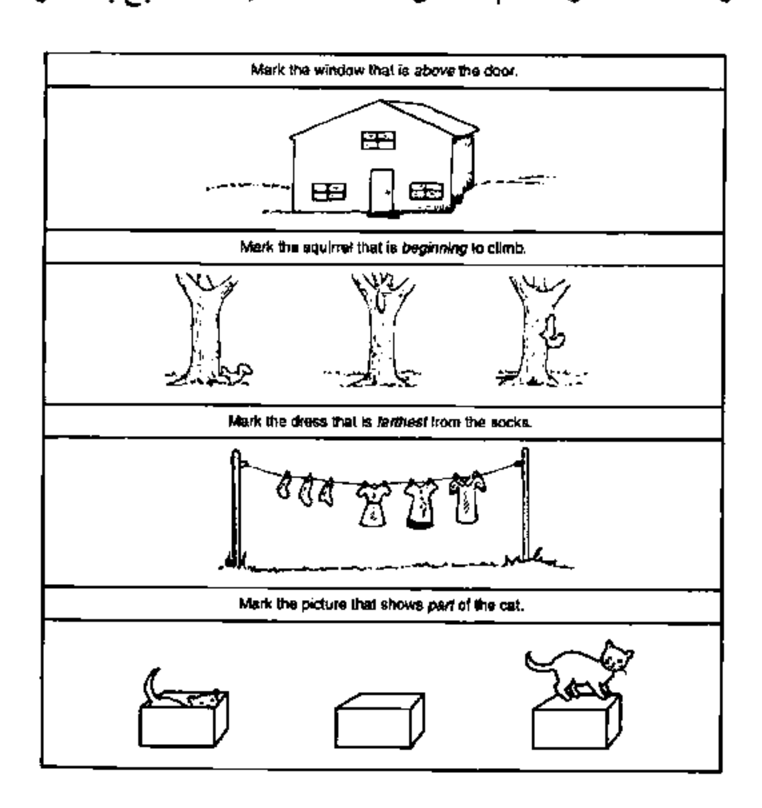
وبينما كانت القدرة على تكوين المفاهيم ضرورية للتطور اللغوي للفرد، إلا أنها لا تكفي لوحدها للتحقيق هذا التطور. إذ على الرغم من قدرة الحيوانات على تعلم بعض المفاهيم البدائية، لا يمكن لهذه الحيوانات التواصل كما يتواصل البشر لسبب أن التواصل الفعال يتطلب القدرة على القراءة والكتابة والحديث باستخدام قواعد النحو والهجاء. وكما هو الحال في القراءة، تتوفر عدد من اختبارات المسح والتشخيص المتعلقة باللغة والمهارات المختلفة المكونة لها. فعلى سبيل المثال يتألف اختبار التشخيص اللغوي MAT6 من اختبارات فرعية تقيس فهم اللغة المسموعة واستخدام علامات الترقيم والأحرف الكبيرة واستخدام اللغة والتعبير الكتابي والتهجئة ومهارات الدراسة. كما تتوفر اختبارات تقيس اللغة المحكية والتهجئة وقدرات الكتابة والحديث وكذلك القدرة على التواصل اللا لغوي (راجع الشكل ٤-٣).

تعد القدرة على كتابة المقالات والتقارير والأبحاث العلمية في نهاية المقررات المختلفة واحدة من المهارات الأساسية المطلوبة من طلاب الجامعات في الولايات المتحدة. ولذلك يتم اختبار عدد كبير من المتقدمين بطلبات قبول في الجامعات الأمريكية في قواعد اللغة الإنجليزية والإنشاء باستخدام اختبارات خدمة الاختبارات التعليمية (كاختبار القبول في صفوف المتقدمين في اللغة

الإنجليزية، واختبارات CELP في المواد المختلفة، واختبارمجلس الكليات SAT II للكتابة). ويتم اختبار الطلاب الذين لا تعد اللغة الإنجليزية لغتهم الأم باستخدام اختبار اللغة الإنجليزية كلغة أجنبية (TOEFL) وهو اختبار يستغرق ثلاث ساعات ويتألف من عدد من الأسئلة المتعددة الاختيارات تم تقسيمها إلى أربعة أقسام تعنى بتقييم استيعاب اللغة المحكية (فهم اللغة المسموعة) وتقييم قدرة الممتَحن على التعرف على الأخطاء النحوية في اللغة الإنجليزية القياسية (بنية اللغة والتعبير) وتقييم استيعاب الطالب للمصطلحات والنصوص التخصصية (المفردات وفهم النصوص).

الشكل ٤-٢

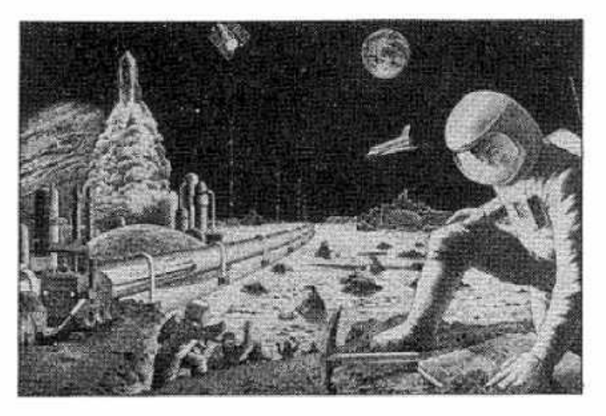
عينة من فقرات اختبار بوهيم المعدل للمفاهيم الأساسية. (حقوق الطبع محفوظة ١٩٨٦ لشركة علم النفس المتحدة. تمت إعادة الطبع بعد الموافقة)



الشكل ٤-٣

صور توضيحية من اختبار اللغة المكتوبة -٣ على الممتَحن أن يؤلف قصة مستوحاة من سلسلة من الصور كتلك المبينة أدناه. (تمت إعادة الطبع بموافقة PRO ED المتحدة، أوستين، تكساس)





اختبارات الرياضيات

تعد مادة الرياضيات من أصعب المواد في المناهج المدرسية والتي ترتفع نسبة الرسوب فيها بشكل كبير. ونظراً لأهمية هذه المادة تتوفر عدد من الاختبارات المسوقة تجارياً التي تقيِّم قدرات الطلاب في هذه المادة الأساسية في بريطانيا والولايات المتحدة وأستراليا وغيرها من البلاد الناطقة باللغة

الإنجليزية. ويضم برنامج F-8 في القرص المرافق للكتاب قوائم بأسماء عدداً من هذه الاختبارات. تقيس بعض الاختبارات المهارات الرياضية في المراحل الابتدائية أو المتوسطة والثانوية أو الدراسة الجامعية أو مقررات ومهن محددة تتطلب مهارات رياضية معينة.

وكما كان الأمر في تقييم اختبار التحصيل في القراءة يمكن تصنيف اختبارات القدرات الرياضية، إلى اختبارات مسح وتشخيص وجاهزية (أو اختبارات تنبؤية). وتغطي اختبارات المسح للقدرات الرياضية العامة مثل اختبارات MAT6 لمسح الإمكانيات الرياضية واختبار ستانفورد للرياضيات في المراحل الدراسية المختلفة من الحضانة إلى الثانوية العامة. كما تقيس اختبارات المسح الرياضية ثلاث نواحٍ هامة في عالم الرياضيات هي المفاهيم العددية والحساب والتطبيقات الرياضية. وتُعنى اختبارات التشخيص للقدرات الرياضية كاختبارات MAT6 التشخيصية واختبار الرياضيات الأساسية المعدل بتقييم النواحي ذاتها التي تقيمها اختبارات المسح إلا أنها تختلف عنها بكونها تحليلية وتقترح حلولاً على عكس اختبارات المسح التي تقتصر على الوصف فقط. وتساعد هذه الاختبارات على تحديد نقاط ضعف وقوة الطالب بغرض تصميم برنامج تعليمي مناسب لهذا الطالب.

ويُستخدم عدد من الاختبارات الرياضية المتقدمة في المراحل التعليمية العليا (أي السنوات الأخيرة من مرحلة الدراسة الثانوية وسنوات الجامعة) كامتحانات CELP للرياضيات العامة وامتحانات CELP في مادة الرياضيات – الجبر والمثلثات والإحصاء... إلخ – واختبارالساعة الواحدة لمجلس الكليات في مجال التحصيل الرياضي في توزيع الطلاب على المستويات الدراسية المختلفة في مادة الرياضيات أو في إعفاء بعض الطلاب من عدد من المقررات أو أعطائهم ساعات معتمدة لمقرارات معينة دون أن يضطروا لحضورحصصها الدرسية. كما يفيد اختبار الرياضيات الذي ينتمي لمجموعة NTE للاختبارات

التخصصية في اختبار المدرسين المتدربين قبل البدء في الخدمة أو العاملين في المدراس والذين يدرسون مادة الرياضيات.

وتساعد اختبارات الجاهزية أو التنبؤية للمهارات الرياضية في التنبؤ بأداء الطلاب في عدد من المقررات الرياضية مثل الجبر للمبتدئين. ويستخدم الاختباران التاليان: اختبار أورلينز – هانا المعدل للتنبؤ بالقدرات في مادة الجبر واختبار أيوا المعدل للقابلية في مادة الجبر، في قياس قدرات الطلاب الذين لم يدرسوا مادة الجبر بعد (وهم طلاب الصفين السابع والثامن). وتتوفر مقاييس لاختبار أورلينز – هانا لطلاب الصفوف ٧ إلى ١٢.

اختبارات العلوم

يعتمد الاتجاه المعاصر في تدريس مادة العلوم في الصفوف المدرسية على تطبيق ما يتعلمه الطالب في هذه المادة على القرارات التي يتخذونها في الحياة العملية. ولهذا تعتمد اختبارات العلوم في مراحل الدراسة الإبتدائية ليس على تذكر الطالب لمعلومات منفصلة سمعها من مدرس العلوم أو قرأها في كتاب المادة بل على التعرف على، ومن ثم محاولة فهم، الأنماط التي تتكرر في المعلومات العلمية. ولا يتوفر عدد كبير من الاختبارات المنشورة التي تعنى بمادة العلوم لطلاب المرحلة الإبتدائية. ومن الاختبارات المتوفرة اختبار بتلر لعلوم الحياة واختبارات فهم العلوم.

يصبح تدريس مادة العلوم أكثر تخصصية في مراحل الدراسة المتوسطة والثانوية إذ يستعاض عن دراسة العلوم بشكل عام بدراسة علم الأحياء والفيزياء والكيمياء وتتبع الاختبارات التحصيلية الوطنية هذا النموذج حيث تستخدم اختبارات العلوم العامة مع طلاب الصفوف ٧ إلى ٩ واختبار الكيمياء العام واختبار الفيزياء العام واختبار علم الأحياء العام لطلاب الصفوف ١٠ إلى ١٦، ويتم اختبار طلاب المدارس الراغبين في الإلتحاق بصفوف المتقدمين أو في

الحصول على عدد من الساعات المعتمدة في مقرر ما من مقررات العلوم باستخدام امتحانات القبول في صفوف المتقدمين في علم الأحياء أو الفيزياء أو الكيمياء أو اختبارات CELP للمادة التي يرغبون في تقييم مستواهم فيها. كما يمكن قبول الطلاب في المقررات الجامعية للعلوم المختلفة بالإعتماد على درجات هؤلاء الطلاب في اختبار مجلس الكليات ذو الساعة الواحدة للتحصيل في مجال علم الأحياءأو الفيزياء أو الكيمياء أو اختبار SAT II للمادة. وتعد الاختبارات التالية اختبارات قيمة في مجال الكيمياء: اختبارات جمعية الكيمياء الأمريكية لطلاب الجامعات واختبارات ACS في الكيمياء العامة والكيمياء العضوية واللاعضوية والتحليلية والفيزيائية والبوليمرية بالإضافة إلى الكيمياء الحيوية. أما في مجال علم الأحياء فينصح باستخدام امتحانات ACT للكفاءة في مجالات التشريح والفيزيولوجية وعلم الأحياء الدقيقة، بينما يطلب من طلاب السنة الأخيرة في الجامعة الذين يرغبون في تدريس مادة علم الأحياء على مستوى المدارس الثانوية التقدم إلى اختبار NTE للمختصين في مجال علم الأحياء.

اختبارات التاريخ والدراسات الاجتماعية

يدرس طلاب المرحلة الإبتدائية مادة التاريخ وغيرها من العلوم الاجتماعية ضمن إطارٍ عامٍ جداً، ولا تتاح لهم فرصة دراسة هذه المواد بالتفصيل إلا في المرحلة الثانوية. ولهذا السبب لا يتوفر عدد كبير من الاختبارات القياسية لمادة التاريخ في سنوات الدراسة الإبتدائية. وقد تم خلال السنوات الأخيرة، إعداد عدد من الاختبارات القياسية للمرحلة الإبتدائية في مادة الاقتصاد وذلك بسبب التركيزالجديد على تدريس مادة الاقتصاد في الصفوف الابتدائية. ومن هذه الاختبارات، اختبار المرحلة الابتدائية لفهم مادة الاقتصاد (للصفين الثاني

والثالث) والاختبار الأساسي في الإقتصاد (للصفوف ٤-٦). ويزداد عدد الاختبارات القياسية عندما ننتقل إلى مرحلة الدراسة المتوسطة إذ تتوفر اختبارات عديدة في مادة التاريخ وغيرها من العلوم الاجتماعية مثل اختبار التحصيل في التاريخ اليهودي واختبار المرحلة المتوسطة في التاريخ الأمريكي والاختبارات السريعة لمقررات التاريخ الأمريكي واختبار المعرفة الإقتصادية. أما في المرحلتين الثانوية ومرحلة السنوات الأولى من الدراسة الجامعية، فيتوفر عدد كبير جداً من الاختبارات في مجالات التاريخ والإقتصاد والعلوم السياسية وعلم النفس وعلم الإجتماع. تُستخدم بطارية اختبارات التحصيل في المواد المختلفة لتوزيع الطلاب على المقررات المختلفة في المرحلة الجامعية ولمنح الطلاب ساعات معتمدة في بعض المواد، ومن هذه الاختبارات اختبار CAT لنهاية مقرر التاريخ الأمريكي واختبار الإقتصاد الاستهلاكي واختبار تاريخ العالم. وتتوفر أعداد أخرى من الاختبارات مثل اختبارات CELP في المواد المختلفة: التاريخ الأمريكي والحضارة الغربية والعلوم الاجتماعية والتاريخ ومبادئ التسويق وعلم النفس للمبتدئين والإقتصاد الكلي للمبتدئين والاقتصاد الجزئي للمبتدئين وعلم الاجتماع للمبتدئين. كما تتوفر كذلك اختبارات الساعة الواحدة لمجلس الكليات واختبارات SAT II في تارخ الولايات المتحدة والعالم والعلوم الاجتماعية. ويطلب من الطلاب الراغبين بالإلتحاق بمقررات الدراسات العليا إجتياز اختبار GRE (اختبار سجل الخريجين) في مواد التاريخ والاقتصاد والعلوم السياسية وعلم النفس وعلم الاجتماع، بينما يستخدم اختبار NTE للمختصين الراغبين في تدريس هذه المواد في المدارس الثانوية.

المهارات الموسيقية والفنية ومهارات استخدام المكتبات

تتوفر، نظراً لتلقي الطلاب دروساً في الموسيقى منذ سنوات الدراسة الأولى، اختبارات قياسية للموسيقى لطلاب الصفين الثالث والرابع

كاختبارات التحصيل الموسيقي واختبارات أيوا للمعرفة الموسيقية واختبارات فارنم الموسيقية. أما اختبارات الموسيقى لطلاب المرحلة الثانوية والجامعية فتضم اختبار نظرية الموسيقى للقبول في صفوف المتقدمين واختبار GRE للموسيقى.

تندر الاختبارات التحصيلية القياسية في مجال الفن لسني الدراسة الإبتدائية بينما تتوفر اختبارات لطلاب الثانوية كإمتحان القبول في صفوف المتقدمين في مادة تاريخ الفن ومادة الرسم. ويستخدم اختبار NTE للمختصين لاختبار الخريجين الراغبين في تدريس مادة الفن في المدارس الثانوية.

كما تتوفر اختبارات أخرى في مجال العلوم الإنسانية مثل اختبار CELP للعلوم الإنسانية. وتتوفر اختبارات تقيس قدرة الطلاب على استخدام المكتبات وغيرها من المهارات التي تساعدهم في دراستهم كاختبار المكتبات ومهارات الدراسة لطلاب الصفوف ٢ إلى ١٢ واختبار مدرسة أوهايو للمكتبات ووسائل الإعلام لطلاب الصفوف ٤ إلى ١٢، وأخيراً، من الجدير بالذكر أن بغض الاختبارات التحصيلية القياسية تقيم مهارات لا ترتبط بمجال التعليم بل بالحياة العملية وعالم الأعمال والمهن.

اختبارات مجال الأعمال والمهن

لا يقتصر استخدام الاختبارات التحصيلية على المدراس والجامعات والمؤسسات التعليمية بل يتعداها ليدخل في مجالات الأعمال والصناعة وبخاصة في اختيار الموظفين أوتوزيعهم على الدوائر المختلفة أو منحهم ترقية أو علاوة أو تراخيص بمزاولة بعض المهن. وتتناول الفصول ٦ إلى ٨ من هذا الكتاب موضوع اختبارات الاستعداد العامة والخاصة بشكلٍ مفصل بينما يحتوي برنامج H-7 الموجود على القرص المرافق لهذا الكتاب على عينات من هذه الاختبارات.

يتم تدريب معظم الحرفيين على أيدي خبراء بعد تعيينهم في وظائف تحتاج لمهارات معينة. وتتألف الاختبارات المهنية من عدد من الأسئلة الشفهية أو الكتابية، كما تعتمد بعض الاختبارات على اختبار المتقدم بطلب التوظيف (بعد أو قبل أن يتم تدريبه) في عينة من المهام التي ستوكل إليه في حال نجاحه في الحصول على الوظيفة. ويمكن أن تقوم دائرة شؤون الموظفين باختبار هؤلاء المتقدمين أو أن يقوم فرد أو مؤسسة متخصصة في مجال الاختبارات بهذه المهمة. ويوضح الشكل ٤-٤ عينة من الاختبارات ذات الأسئلة متعددة الخيارات لعدد من الحرف توزعها دار النشر ماكغرو - هيل / لندن.

ويتضمن البرنامج الوطني للكفاءة المهنية عدداً كبيراً من الاختبارات التي طورتها وأعدتها خدمة الاختبارات التعليمية ETS. وتتوفر اختبارات للمهن التالية: صيانة السيارات وتقديم المشروبات في الحانات والتجميل وصرف العدسات اللاصقة والإطفاء وخدمات الأجانب والتأمين والشرطة والغولف للمحترفين وغيرها. ولا تعتمد كل هذه الاختبارات على أسئلة الخيارات المتعددة أو على الورقة وقلم الرصاص فقط، فمن الممكن مقابلة الممتحنين أو أن يطلب منهم كتابة رد على مذكرة ما أو المشاركة في مناقشة مشكلة أو قضية ما أو أن يقوموا بمهمة ما يتم تقييم أدائهم فيها لاحقاً.

على الرغم من قيام أجهزة الحاسب بعدد هائل من المهام المكتبية، لا تزال الشركات والمؤسسات المختلفة بحاجة إلى الكَتَبَة وبأعداد كبيرة. وتتوفر اختبارات تحصيلية واختبارات للاستعداد تقيس قدرة الأفراد على تعلم وتأدية عدد من المهام المكتبية كاختبار مهارات المكتب التحصيلي. ويوضح الشكل ٤-٥ عدداً من الفقرات الاختبارية التي تدخل في الاختبارات الفرعية المكونة لهذا الاختبار والتي تقيم المهارات المطلوبة لعدد كبير من الوظائف المكتبية. ويوجد بالإضافة إلى الاختبارات الفرعية الأحدى عشر التي يوضحها الشكل ٤-٥ اختبار فرعي آخر هو اختبار التعليمات الشفهية والذي يطلب من الممتحنين الاستماع إلى شريط تسجيلي ثم الإجابة على الأسئلة المرتبطة به.

كما تتوفر عدد من الاختبارات لمهارات المكتب ومهارات الحاسب بالإضافة إلى المهارات الأساسية كالقراءة والحساب والقدرة اللغوية (كما توضح قائمة اختبارات عالم الأعمال والحرف التي يضمها برنامج (F-8). كما يضم برنامج F-8 قائمة أخرى بأسماء الاختبارات التحصيلية التي يتوجب على عدد من الأفراد في مجالات التدريس أو المحاسبة أو الموارد البشرية اجتيازها بغرض الحصول على ترخيص بمزاولة المهنة. ومن الأمثلة البارزة عن هذه الاختبارات امتحانات المدرسين الوطنية (NTE) في عدد كبيرمن المقررات المدرسية. وغالباً ما لا تعتمد اختبارات بعض المهن كالطب والمحاماة على اختبارات تنشر على مستوى الدولة بل تقوم الولايات بتطوير وإجراء اختبارات المجلس الطبي للولاية التي يرغب المتقدم بممارسة المهنة فيها.

الشكل ٤-٤

عينة لفقرات اختبار من عدد من اختبارات "الورقة وقلم الرصاص" للمعرفة الحرفية (تمت إعادة النشر بموافقة د. رونالد ت. رامزي)

اختبار تقنيي الكهرباء

١- يقوم جهاز التحكم الحركي التواقتي بتشغيل المحرك كمحرك حتّى ثم:

أ- يقوم بمواقتة حركته باستخدام حث التيار المباشر

ب- يقوم بتشغيله كمحرك تيار مباشر متزامن

ج- يقوم بمواقتة المولد الكهربائي مع تردد التيار

د- يقوم بمواقتة مولد الطاقة مع التردد

اختبار تقنيي الميكانيك

٢- يتم وضع منفذ للهواء في خزان السوائل لكي:

أ- يسمح بتدفق الهواء

ب- يبرد السائل

ج- يبرد المخزن

د- يتحكم بتلوث الهواء

اختبار السباكين وعمال التمديدات الصحية

٣- مفاتيح البراغي ذات النهايات الدائرية المغلقة آمنة أكثر من المفاتيح ذات النهايات المفتوحة بسبب:

أ- كونها أكبر حجماً

ب- كونها أصغر حجماً

ج- كون احتمال إنزلاقها أقل

د- كونها أقدر على تحريك البراغي

اختبار عمال صيانة الأبنية

٤- ما هي الأداة الأنسب لضرب المسامير؟

أ- المطرقة ذات النهاية المشطورة

ب- المطرقة ذات النهاية الدائرية

ج- المطرقة الثقيلة

د- المطرقة المطاطية

اختبار عمال اللحامة

٥- ما لون خرطوم الأوكسجين؟

أ- أخضر

ب- أصفر

ج- أسود

د- أحمر

الشكل ٤-٥

عينة من اختبار المهارات المكتبية (أعيد طبعه بموافقة ماكغرو- هيل / لندن)

اختبار مهارات التدقيق

القائمة الصحيحة

ألكساندر، توم	٧٦١	٣٢٣	٩٦٤
دايفيز، بيني	٨٩٠	٦٤٧	٣٨٥
هس، لاري	٢٧٦	٠٦٢	٥٩٦
سانتوس، دولرس	٢٧٥	١٨٦	٢٣٣
ويلسن، ريتشارد	٢٤٩	١٨٧	٦٠٢

القائمة المطلوب تدقيقها				صح	خطأ
١- دولورس سانتوس	٢٧٥	١٨٦	٢٢٣	□	■
٢- توم ألكساندر	٧٦١	٣٢٣	٩٦٤	■	□
٣- ريتشارد ويلسن	٢٤٩	١٨٧	٦٠٢	□	■

اختبار وضع الرموز

مفتاح الرموز

٣٤ ذكر

٢١ أنثى

M بالغ

U مراهق

Z طفل

يتبع كل فقرة دائرة تحتوي على خمسة رموز محتملة. إن مهمتك هي التعرّف على الترتيب المناسب للرموز وتضع إشارة "X" في الدائرة المناسبة، لقد تم وضع إشارة "X" على الإجابة الصحيحة للمثال رقم ١.

الأمثلة:

١- أنثى يافعة (34U) (34M) (86Z) (21M ✗) (21U)

٢- طفل ذكر (12Z) (34Z ✗) (34U) (21U) (34M)

اختبار تنظيم الملفات

انظر إلى الأسماء في عمود الملفات المراد تنظيمها ثم ابحث عن رقم الملف الذي يمكن تنظيم المعلومات ضمنه من بين الملفات الموجودة وضع إشارة x إلى جانب الخيار الصحيح، وإذا لم يتوفر رقم في مكان الخيار الصحيح فضع إشارة "X" في الدائرة الفارغة.

الملفات الموجودة

١- فيليب جنكنز

٢- ج. ك. كايل

٣- شركة توماس موريس المتحدة

٤- شركة بولسون المتحدة

٥- سالي وايت

الملفات المراد تنظيمها

أ . ب . رينولدز ① ② ③ ④✗ ○

جون جونز ①✗ ② ③ ④ ○

اختبار المهارات العددية

الأمثلة:

١ . ٦١ + ٢٤

□ ٧٠

□ ٧٤

□ ٧٥

□ ٨٤

■ ٨٥

٢ . ١/٤ - ٣/٤

□ ٢

□ ١

■ ١/٢

□ ١/٤

□ ٣/١٦

٣- يكلف ١٠٠٠ مغلف صمم خصيصاً لشركة ما ٧٩,٠٠ دولاراً. ما كلفة ١٠٠٠٠ مغلف؟

□ ٠,٧٩ دولاراً

□ ٧,٩٠ دولاراً

□ ٧٩,٠٠ دولاراً

■ ٧٩٠,٠٠ دولاراً

□ ٧٩٠٠,٠٠ دولاراً

اختبار ملء الاستمارات

أبلغت شركة صوت اليوم في الساعة ٨:٣٠ من صباح يوم الخامس عشر من أوكتوبر عام ١٩٧٧ عن سرقة في مقرها الكائن في ٣٩٠٧ في بالم أفنيو، ويستا (كاليفورنيا)، هاتف:٦٨٩-٧٧٣٤، وتم الإبلاغ عن سرقة آلتي تسجيل وجهازي تكبير الصوت وعلبتي أشرطة كاسيت فارغة. وقد تم إلحاق الضرر بالباب الخلفي للدخول إلى مقر الشركة. كما يعتقد أنه تم إفتعال حريق في مبنى مجاور لإستبعاد الشك في السرقة. رقم البلاغ هو 789A3.

املأ الفراغات بالمعلومات المناسبة

رقم البلاغ ....

التاريخ: .... وقت البلاغ: ....

اسم الضحية (أو الشركة): ....

مكان الحادث: .... هاتف:..............................

حجم الأضرار: ....

اختبار القواعد

ألبرت .... الرسائل التي أمليتها عليه.

☐ يطبع

☐ تم طبع

☐ يطابع

■ طبع

☐ تطبع

اختبارات علامات الترقيم

يختبر هذا الاختبار معرفتك بعلامات الترقيم واستخدام الحروف الكبيرة. يتألف الاختبار من جمل تم تقسيمها إلى أربعة أقسام وقد تحتوي كل جملة على عدد من الأخطاء أو قد لا تحتوي على أية أخطاء. ضع إشارة x إلى جانب العبارات التي تحتوي على أخطاء أو إلى جانب العبارة التي تقول "لا أخطاء".

أمثلة:

١- سألتني السيدة براون (١) | إذا كنت قد حضرت (٢ ✗) | التقارير الأسبوعية (٣) | لاجتماع يوم الجمعة (٤) | لا يوجد خطأ (٥) |

٢- نهركانتون (١ ✗) | نادي جلي سيقدم (٢) | عرضاً في الجامعة (٣) | السبت القادم (٤) | لا يوجد خطأ (٥) |

اختبار فهم النصوص

طلب من مارك إصلاح آلات التصوير في الطابقين الرابع والسابع قبل القيام بأعمال الصيانية الروتينية التي يقوم بها يومياً. إلا أنه يجب عليه إصلاح الإضاءة في الطابق الثالث أولاً إذ تم تلقي عدد من الشكاوى عن هذا الأمر.

ما الذي يتوجب على مارك فعله أولاً:

☐ إصلاح آلة التصوير في الطابق الرابع

☐ إصلاح آلة التصوير في الطابق السابع

☐ إجراء عمليات الإصلاح الروتينية

■ إصلاح الإضاءة في الطابق الثالث

☐ إصلاح الإضاءة في الطابق الرابع

اختبار التهجئة

الرجاء إرسال....... قبل يوم الجمعة.

□ اللرد

□ الرض

□ الردد

□ الراد

■ الرد

اختبار الطباعة

اختفت رومانا وودستوك التي تعودت أن تعود إلى المنزل قبل الساعة الحادية عشرة. أتصلت بعائلتها الساعةالثانية ولم تعرف أخبارها بعد ذلك الوقت. لا تزال حتى الآن خارج المنزل[١].

اختبار المفردات

النسبة المئوية

□ كامل

□ متأخر

□ مجموعة

■ حصة

□ مجموع

(١) يكون النص مكتوباً بخط اليد ويطلب من الممتحن طباعته باستخدام الكمبيوتر أو معالج الكلمات. (ملاحظة المترجم).

التدقيق

- □ التسلق
- □ الجمع
- ■ الفحص
- □ النشر
- □ البريد

الخلاصة

استخدمت الاختبارات التحصيلية الكتابية للمرة الأولى في القرن الثالث عشر، إلا أنه لم يتم استخدام اختبارات التحصيل القياسية والتي تتألف من أسئلة متعددة الخيارات إلا في بداية القرن العشرين. وقد أُتْبِعَ نشر الاختبارات القياسية في مجالات المهارات الأساسية كالحساب والتهجئة والخط، نشر بطاريات التحصيل (كاختبار ستانفورد للتحصيل وامتحانات أيوا لمحتوى التعليم الثانوي وغيرها) في العشرينيات من القرن الفائت. وتعد الاختبارات التحصيلية أكثر أنواع الاختبارات التعليمية والنفسية استخداماً، إذ تبيع الشركات الأمريكية المسوقة للاختبارات ما قيمته ٢٠٠ مليون دولار من الاختبارات التحصيلية المدرسية كل عام كما تبيع شركات في بريطانيا وأستراليا وجنوب أفريقية وكندا أعداداً لا بأس بها من هذه الاختبارات.

ولا تقتصر فائدة درجات اختبارات التحصيل على تحسين وتطوير العملية التعليمية فحسب بل تساهم كذلك في توزيع وتصنيف واختيار الأفراد في مجالات التعليم والتوظيف. ويفيد أداء عدد كبير من الطلاب في هذه الاختبارات كذلك في تقييم المدرسين والمناهج والبرامج والسياسات التعليمية.

ويتم استخدام الاختبارات التحصيلية القياسية واختبارات التحصيل التي يضعها المدرسون على حدٍ سواء. إلا أن استخدامات كلا النوعين تتنوع وتختلف بعض الشيء، ففي حين تركز الاختبارات التي يضعها المدرسون على الأهداف التعليمية التي تم تغطيتها في حصص هذا المدرس، تهدف الاختبارات التحصيلية القياسية إلى قياس المهارات والقدرات العامة (كالقراءة والحساب والمهارات اللغوية) التي تتشابه في عدد كبير من المدارس. ويفيد كلا النوعان في عمليات التقييم التكويني أو التراكمي وفي خدمة أهداف المسح (أو التقييم العام) والتشخيص والتنبؤ بالأداء، إلا أن الاختبارات القياسية تكون عادة أكثر ثباتاً ويمكن مقارنة أداء الطالب فيها مع معايير ثابتة ترافق الاختبار.

نموذجياً؛ تبدأ المدارس باختبار طلابها مستخدمة البطاريات التحصيلية كل عام أو عامين في المرحلة الإبتدائية وكذلك في المرحلة المتوسطة. أما للحصول على صورة أكثر وضوحاً عن أداء الطالب في القراءة أو المواد الأدبية أو الرياضيات أو غيرها من المواد المختلفة، فيتم اختبار الطالب باستخدام الاختبارات التحصيلية التخصصية. وتتوفر اختبارات المسح لجميع المقررات المدرسية أما اختبارات التشخيص فلا تتوفر إلا للقراءة والتهجئة والرياضيات. ويمكن الحصول على عدد كبير من اختبارات التحصيل للمراحل الثانوية والجامعية من خدمة الاختبارات التعليمية ETS لاستخدامها في مجالات قبول الطلاب أو إعفائهم من مقررات معينة أو منحهم عدد معين من الساعات المعتمدة.

كما يبيع ناشرو الاختبارات عدداً من الاختبارات التحصيلية في مجالات: القراءة (اختبارات مسح وتشخيص وجاهزية) واللغة الإنجليزية (تقنيات اللغة والإنشاء) واللغات الأجنبية والرياضيات والعلوم والتاريخ والفنون الجميلة والعلوم الاجتماعية والإنسانية وغيرها. ولا يقتصر استخدام بطاريات التحصيل والتي تختبر مادة واحدة على مجالات التعليم إذ تستخدم مؤسسات الدولة والشركات

الصناعية والتجارية والقوات المسلحة وغيرها هذه الاختبارات عند إتخاذ قرارات تتعلق بأمور التوظيف. ويمكن الحصول على قائمة بأسماء بطاريات التحصيل وناشريها والعمر الذي صممت من أجله عن طريق برنامج F-8 الموجود على القرص المرافق لهذا الكتاب. كما يمكن الحصول على معلومات أكثر دقة عن هذه الاختبارات بالعودة إلى كتابي: الاختبارات المنشورة ٢ والكتب السنوية للمقاييس العقلية.

مصطلحات للمراجعة

راجع معاني المصطلحات المدرجة أدناه والتي تم استخدامها في هذا الفصل. الرجاء مراجعة فهرس التعريفات أو المعجم في حال وجود أي التباس حول معاني المصطلحات.

اختبار تحصيلي	اختبار الجاهزية لتعلم القراءة
بطارية اختبارات التحصيل	الاختبارات القياسية
الاختبارات ذات المرجعية المحكية	التقييم التراكمي
الاختبارالتشخيصي	اختبار المسح
التقييم التكويني	الاختبار الذي يضعه المدرس
الاختبارات ذات المرجعية المعيارية	اختبارات مجال الأعمال
اختبارات التنبؤ بالأداء	اختبار عينة المهام الوظيفية للمتقدم بطلب التوظيف

المراجع المقترحة

Green, K. E. (1991). Teachers and standardized testing of students. In K. E. Green (Ed.), *Educational testing: Issues and applications* (pp. 197–224). New York: Garland.

Hambleton, R. J., & Jurgensen, C. (1990). Criterion-referenced assessment of school achievement. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence and achievement* (pp. 456–476). New York: Guilford Press.

Katz, L. J., & Slomka, G. T. (1990). Achievement testing. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (2nd ed., pp. 123–147). New York: Pergamon Press.

Koretz, D. (1992). State and national assessments. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed.) (Vol. 4, pp. 1262–1267). New York: Macmillan.

Linn, R. L. (1992). Achievement testing. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed.) (Vol. 1, pp. 1–11). New York: Macmillan.

Madson, D. J. (1991). Legal issues in standardized achievement testing. In K. E. Green (Ed.), *Educational testing: Issues and applications* (pp. 225–243). New York: Garland.

Nitko, A. J., & Lane, S. (1990). Standardized multilevel survey achievement batteries. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence and achievement* (pp. 405–434). New York: Guilford Press.

الملحق ٤-١

تقييم نقدي لبطارية اختبارات تحصيلية قياسية

من كتاب اختبارمستويات التحصيل للمناهج متعددة المستويات (SCALE) الذي نشرته مؤسسة الخدمات النفسية في المنطقة الغربية في لوس أنجلوس بولاية كاليفورنيا عام ١٩٩٢، والكتاب من تأليف ف. و. دوهرتي و ج. هـ. رويد.

وصف عام للاختبار

تم تصميم الاختبارات التي تتألف منها هذه البطارية بهدف قياس التحصيل على مستوى الأفراد والمجموعات من طلاب الصفوف الثالث إلى الثامن في الحسابات الرياضية والقراءة واستخدام اللغة. ويتألف الاختبار لكل من المواد الثلاثة الآنفة الذكر من كتيب يحتوي على أسئلة خيارات متعددة مقسمة إلى تسع مستويات لمدى الصعوبة. وقد تم تصميم ووضع هذه الاختبارات وقياسها كجزء من برنامج لتطوير المناهج والاختبارات قامت به مدارس الدولة في بورتلاند في ولاية أوريغون وذلك بغرض استخدام هذه الاختبارات في مجالات قبول وتوزيع الطلاب في المدارس والفصول المختلفة وفي تشخيص المشاكل التعليمية لعدد من الطلاب. وقد تم اتباع، في تصميم بطارية SCALE، إجراءات مشابهة لتلك المستخدمة في تصميم اختبارات مستويات التحصيل في بورتلاند والتي تعد أكثر شمولا من SCALE.

تتنوع أشكال ومحتويات الأسئلة الثلاثة والعشرين التي يتألف منها كل مستوى صعوبة في اختبارات SCALE باختلاف مستوى الصعوبة والأهداف

التعليمية المرتبطة بالمنهج التي يتوجب على الاختبار تقييمها. فعلى سبيل المثال، يتوجب على اختبارات الحسابات الرياضية أن تقيم أداء الطلاب في مجالات الطرح والجمع والقسمة والضرب والأعداد العشرية والكسور والنسب المئوية، بينما تقيس اختبارات القراءة تحصيل الطالب من حيث معرفة معاني الكلمات والفهم الحرفي للنصوص والفهم التحليلي (التأويلي) لمعاني النصوص بالإضافة إلى الفهم التقييمي للنصوص. أما اختبارات استخدام اللغة فتعنى بتقييم عملية إنشاء (تأليف) النصوص وبنية النص والاستخدام السليم للقواعد وعلامات التقييم والأحرف الكبيرة.

تستغرق الاختبارات الثلاثة ٧٩ دقيقة موزعة على النحو التالي: ٢٠ دقيقة لاختبار الحسابات الرياضية و٢٠ دقيقة لاختبار القراءة و١٥ دقيقة لاختبار استخدام اللغة بالإضافة إلى ٢٤ دقيقة ترصد لقراءة التعليمات. ويساعد كون الاختبارات مرتبطة بشكل مباشر بمستويات المنهج الذي يدرسه الطلاب على اختبار الطالب وفقاً لمستوى قدرته. ويبدأ الاختبار بالمستوى الذي تم تقديره على أنه المستوى المناسب لهذا الطالب. ويمكن تحديد هذا المستوى عن طريق إدخال مجموعة من المعلومات عن الطالب في الجدول الموجود في الملحق (أ) كالسنة الدراسية التي ينتمي إليها الطالب ومعدل درجاته في الاختبارات أو عمليات التقييم السابقة التي خضع لها. وتساعد أوراق الإجابة التي تسمح بالتصحيح الذاتي السريع (وضع أوتوماتيكي للدرجات) للإجابات (خلال ٣ أو ٤ دقائق لكل طالب وللبطارية بأكملها) في إعادة اختبار أي طالب كان معدل درجاته عالياً أو متدنياً بشكل ينفي عن الاختبار صفة الصدق.

تستند اختبارات SCALE إلى نماذج راش Rasch من نظرية الإستجابة إلى الفقرة، وقد تمت معايرة الأسئلة وفقاً لمقياس لصعوبة الأسئلة وقدرات المتحَنين وتم، بناءً على هذه المعايرة، تحديد الحدود العامة لمعدلات درجات الاختبار بين

١٤٠ و٢٦٠ بوسط حسابي ٢٠٠ وإنحراف معياري ١٠، وقد ساهم استخدام نظرية الاستجابة إلى الفقرة في تطوير اختبارات SCALE إلى رفع ثبات هذه الاختبارات أعلى من المستوى المتوقع للاختبارات القصيرة المشابهة.

وتتم مقارنة درجات الطلاب في بطارية SCALE بالمقاييس المبنية على المنهج ومعايير السنة الدراسية التي ينتمي إليها الطلاب (الدرجات القياسية والمعايير المئوية ومعايير NCE). وتتوفر هذه المعلومات في الملحق (ب) الموجود في دليل الاستخدام المرافق للاختبار. ويفيد المقياس المبني على المنهج التعليمي في التنبؤ بإحتمالية إتقان الطلاب للمهام التي يطلب المنهج منهم القيام بها. وتفيد المعلومات الناتجة عن هذه التنبؤات في تقسيم الطلاب إلى مجموعات أو في تخطيط المناهج أو في توزيع الطلاب على المستويات والمقررات المختلفة. ويحتوي الجانب الخلفي لورقة الإجابة لكل اختبار على درجة صعوبة الفقرات الاختبارية تبعاً لمؤشر راش لصعوبة الفقرات. وتتوفر هذه المعلومات في الملحق (ج) من دليل المستخدم والذي يحتوي كذلك على معلومات عن ارتباط كل فقرة اختبارية بنواحي المنهج المتعددة. ويقوم واضع الدرجات بعد تحديد موقع درجات الممتحن في اختبار SCALE على الجانب الخلفي لورقة الإجابة برسم خط يمثل مستوى درجات الطالب (أو مستوى الإحباط أي المستوى الذي تؤدي صعوبته إلى إصابة الطالب بالإحباط) والذي يتنبأ بإحتمالية ٥٠٪ لإجابة الطالب بشكل صحيح على الأسئلة ضمن هذا المستوى.

كما يمكن رسم خطين آخرين على ورقة الإجابة: أحدهما على مقياس SCALE ناقص الرقم ١٠ (مستوى التدريس)، ويمثل هذا الخط احتمالية ٧٥٪ لإجابة الطالب بشكل صحيح على الأسئلة ضمن هذا المستوى؛ أما الط الآخر وهو خط الإتقان Mastery Level فيمكن رسمه عند نقطة مساوية لدرجات الطالب على مقياس SCALE ناقص الرقم ٢٠ ويعطي هذا الخط إحتمالية ٩٠٪ لإجابة الطالب بشكل صحيح على الأسئلة ضمن هذا المستوى.

ويحتوي دليل المستخدم لاختبارات SCALE على معلومات تفصيلية عن الخصائص السايكومترية للاختبارات. ويتراوح الثبات التقليدي (الإتساق الداخلي) للاختبار بين ٠,٨٠ و ٠,٨٧ لاختبار الحسابات الرياضية و٠,٦٨–٠,٨٧ لاختبارات القراءة و٠,٧٢–٠,٨٣ لاختبارات استخدام اللغة. كما يحتوي دليل المستخدم على تقديرات لثبات الاختبار مبنية على نظرية IRT وعلى الخطأ المعياري للقياس وعلى استخدامات المعلومات الناتجة عن الاختبار. كما يضم الدليل اثباتات على صدق اختبارات SCALE (أي قدرتها على قياس النواحي التي تهدف إلى قياسها) من حيث ارتباطها بالمناهج التدريسية ودقة معاييرها في التنبؤ بقدرات الطلاب وإرتباط أقسامها الثلاثة في طريقة قياس العمليات العقلية. كما يحتوي الفصل الخامس من دليل المستخدم على معلومات سيكومترية أخرى تتضمن دليلاً من عمليات المعالجة للصعوبات التعليمية وفي دراسات التغير والنمو عند الطلاب وغيرها من المعلومات المرتبطة بفقرات الاختبار (أحادية البعد، ومتسقة التمييز بين الفقرات، وتقع على المنحنى التشخيصي للفقرات، وغيرها من الفقرات الإحصائية).

النقد

يتألف دليل المستخدم لبطارية SCALE من ١٣٥ صفحة ويحتوي على معلومات تفصيلية عن تصميم الاختبار وطريقة وضع ومقارنة درجاته بالإضافة إلى الخصائص السايكومترية للاختبار. عند معاينة الاختبار للمرة الأولى، يبدو وكأنه نسخة اختبار جماعية عن الاختبار التكيفي تنقصها درجة التحكم العالية في الاختبار التي تميز الاختبارات التكيفية الفردية المحوسبة. يُعطى كل ممتحن ورقة إجابة واحدة وكتيب للأسئلة ويتوقع من الممتحَنين البدء بالأسئلة المناسبة لمستواهم مما يمكن أن يؤدي إلى صعوبات في التواصل مع الممتحَنين فيما يتعلق بأين يبدأ الاختبار وأين ينتهي بالنسبة إلى كل طالب. وعلى الرغم من أن ورقة

الإجابة توضع عند نقطة البدء لكل ممتحَن، يبقى احتمال وقوع الخطأ كبيراً. كما يعد إطلاع الممتحن على جميع الأسئلة التي يتكون منها الاختبار (ليست الأسئلة المتعلقة بالاختبار الحالي فقط) عائقاً في وجه دقة إعادة اختبار الطالب في مستوى مختلف.

والمشكلة الأخرى المقلقة والتي يعاني منها الاختبار كون تعليماته معقدة جداً مما قد يؤثر على أداء الطلاب الأقل مقدرة. كما تعد طريقة تحديد المستوى المناسب لاختبار الطالب طريقة عشوائية وتنقصها الدقة. ففي حالة الاختبارات التكيفية المحوسبة يعطى الطالب سؤالاً أقل صعوبة في حال إجابته بشكل خاطئ على السؤال السابق أو سؤالاً أكثر صعوبة في حال إجابته على السؤال السابق بشكل صحيح، أما في اختبارات SCALE فيتوجب على الطالب الإجابة على جميع الأسئلة التي تقع ضمن المستوى الذي تم تحديده له. ولا يمكن معرفة مدى مناسبة هذا المستوى للطالب إلا بعد إنتهائه من الإجابة على جميع الأسئلة وبعد فراغ القائم على الاختبار من وضع الدرجات. بعد هذا كله يمكن إعادة اختبار الطالب ضمن مستوى أعلى أو أدنى تبعاً لنتيجته. ولهذا السبب ترتفع إحتمالية إعادة الاختبار وهو أمر لم يتعرض له دليل المستخدم بالشكل المطلوب. وأعتقد أن التجارب التي أجريت على الاختبار لم تكشف عن أية مشكلات في عملية إعادة الاختبار، إلا أنني، ومع الأسف، لست على ثقة من صحة إعتقادي هذا.

لعل من أهم المشكلات الأخرى التي يعاني منها الاختبار كون الإجابات مطبوعة خلف ورقة الإجابة مما يسمح للعديد من الطلاب بالغش. وعلى الرغم من أن الهدف من وضع الإجابات على ورقة الإجابة هو تسهيل عمل المصحح، يبقى السؤال الهام هنا هو: كم من الطلاب سوف يغش وكيف سيعرف المصححون إذا وقع الغش أم لم يقع؟

يرى مصممو SCALE بطارية اختباراتهم هذه كأداة قياس أو اختيار أولية وليس كمقياس شامل للتحصيل المعرفي. ويعبر الناشر عن هذه الفكرة بقوله: "إن بطارية SCALE هي الخيار المثالي لكل الراغبين في الحصول على تقييم سريع ودقيق لمستويات التحصيل لدى الطلاب." إلا أن السرعة قد تؤثر على الدقة فيتم التضحية بالشمولية لزيادة كفاءة الاختبار.

ومن الجدير بالذكر أن العناية والدقة التقنية التي تم من خلالها إعداد وتطوير بطاريات اختبارات SCALE بالإضافة إلى المتابعة بالبحث والدراسات لعمل هذه الاختبارات يعد أمراً جديراً بالثناء. وعلى الرغم من أن معظم العاملين في مجال القياس النفسي يفضلون نماذج IRT ذات المؤشرين على نموذج راش ذو المؤشر الواحد، يبدو أن نموذج راش قد خدم مصممي بطارية SCALE بشكل جيد. إلا أنه يبقى لدي بعض التحفظات على هذه البطارية ولذلك فإنني أنصح المدارس بتجربة الاختبارات على عينة من طلابها ودراسة التغذية الراجعة عن مدى خدمة هذه البطارية لأغراض المدرسة قبل الإلتزام بشرائها. كما يجب أن يؤخذ بعين الاعتبار ضرورة تدريب المدرسين وغيرهم من مستخدمي هذه البطارية على طريقة إجراء ووضع درجات الاختبارات والتي تعد أكثر تعقيداً من طرق إجراء وتصحيح الاختبارات غير التكيفية.

٥

الاختبارات التحصيلية الشفهية والأدائية

يعتبر الاختبار ذو أسئلة الخيارات المتعددة والذي نشرت أول نسخة منه (اختبار كنساس للقراءة الصامتة) في عام ١٩١٥ أحد أهم مساهمات علم النفس في الحضارة الإنسانية في القرن العشرين. ويرى واضعو الاختبارات التحصيلية ذات أسئلة الخيارات المتعددة هذه الاختبارات كأدوات قياس فعالة وقادرة على تقييم عدد كبير من الأهداف التعليمية والمهنية. وتقيس هذه الاختبارات في أشكالها القياسية تحصيل الأفراد الأكاديمي ومقدار إنتاجيتهم بالمقارنة مع المعايير المقبولة أو مع أداء عينة من الأفراد تمثل مجموع السكان.

وعلى الرغم من الميزات العديدة التي تتمتع بها الاختبارات التحصيلية القياسية واختبارات القبول الجامعي وغيرها من اختبارات الخيارات المتعددة، تعرض هذا النوع من الاختبارات إلى الإنتقاد الشديد منذ اليوم الأول من استخدامه. ومن ضمن هذه الانتقادات: تركز الاختبارات ذات أسئلة الخيارات المتعددة على الإجابة وليس على عمق التفكير الذي قاد إلى هذه الإجابة أو على المهارة التي تم التعبير عن الإجابة من خلالها، كذلك يفضل هذا النوع من

الاختبارات الطلاب الحاذقين سريعي البديهة والقادرين على القراءة بسرعة على الطلاب المبدعين والهادئين والذين يفكرون بالأمور بعمق وروية (بلاك، ١٩٦٢؛ هوفمان، ١٩٦٢). ومما لا شك فيه أن تركيز الاختبارات التحصيلية على التعرف على الإجابة الصحيحة بدلاً من تذكرها وعلى حفظ المعلومات بدلاً من التحليل والتفكير والصياغة له تأثير سلبي على العملية التعليمية وعلى التعرف على مقدرة الأفراد الحقيقية. وهذا الأمر هو من أهم التحفظات على الاختبارات ذات أسئلة الخيارات المتعددة. كما يرى نقاد الاختبارات التحصيلية ذات أسئلة الخيارات المتعددة أن هذه الاختبارات هي مقاييس سيئة للقدرة والتحصيل وتؤدي إلى تدني مستوى التعليم وتشجع الطلاب على تبني عادات دراسية سيئة.

ولم يسكت مناصرو الاختبارات ذات الأسئلة متعددة الاختيارات في وجه هذه الإنتقادات اللاذعة إذ عبروا عن إستيائهم من هذه الآراء التي اعتبروها آراءاً خاطئة أو تعميمات غير مقبولة. إلا أنهم يعترفون بأن هذا النوع من الاختبارات يعاني من المحدودية ويغفل تماماً تطويرمهارات الكتابة والقراءة عند الطلاب. ويعد استخدام الاختبارات المقالية التي تتطلب أن يبرر الطلاب إجابتهم وأن يدعموها بالأدلة والبراهين حلاً منطقياً لهذه المشكلة (ديفيد، ١٩٨١) إذ أن إضافة أسئلة مقالية للاختبارات وتقييم إجابات الطلاب من حيث الشكل والمضمون ثم إعطاء الطلاب معلومات عن طريقة تحسين أدائهم يمكن أن يفيد في تدعيم قدرة الطلاب في التعبير عن أنفسهم.

وقد أدى الاهتمام بمهارة الكتابة في السنوات الأخيرة إلى نشر عدد كبير من الاختبارات القياسية التي تغطي معظم مراحل الدراسةالمدرسية (كاختبار CAP لتقييم الكتابة وبرنامج ستانفورد لتقييم الكتابة واختبار عملية الكتابة) والمرحلة الجامعية (كاختبارالساعة الواحدة لمجلس الكليات في مجال الكتابة

واختبار الكتابة SAT II). كما أدى الاهتمام بالكتابة أيضاً إلى إضافة اختبارات كتابة اختيارية على بطاريات اختبارات التحصيل التي لم تحتو أصلاً على اختبار للكتابة (كاختبار MAT7 للكتابة وتقييم أيوا لمهارات الكتابة). وتصف الفئة الثالثة من برنامج F-8 أدوات قياس قدرات الكتابة التي تقيم قدرة الطلاب على توليد الأفكار وتنظيمها والتعبير عنها بدلاً من الإكتفاء بالإشارة إلى معرفتهم بقواعد اللغة الإنجليزية وميكانيكيتها.

إن ردود الفعل على الادعاء بأن استخدام الاختبارات الموضوعية ذات أسئلة الخيارات المتعددة بشكل أساسي من الممكن أن يؤدي إلى تدني مستويات التدريس وتحول عملية التقييم إلى عملية سطحية قادت إلى إنعاش الاهتمام بالكتابة وغيرها من المهارات الذهنية، إذ بذل العاملون في مجال التقييم جهوداً جبارة لتطويرأدوات قياس حقيقية تركز على مهارات التفكير العليا لدى الطالب كالتحليل والإستدلال بالإضافة إلى التعبير الشفهي والكتابي عن الأفكار. وفي هذا النوع من أدوات التقييم المفتوحة، يتم منح الدرجات للطلاب ليس على إجاباتهم الصحيحة فحسب بل على الطريقة التي توصلوا من خلالها إلى الإجابة. وعلى عكس الاختبارات التقليدية التي يعمل كل طالب فيها بمعزل عن الآخرين، تتميز اختبارات الأداء هذه بأنها تسمح لمجموعة من الطلاب بالعمل معاً على إجراء تجربة ما وتحليل نتائجها أو على التوصل لإجابة عن سؤالٍ ما أو على إنتاج شيء ما.

يتناول القسم الأول من هذا الفصل موضوع الاختبارات الشفهية والتي تعد نوعاً من اختبارات تقييم الأداء، بينما يركز القسم الثاني على أنواع أخرى من اختبارات تقييم الأداء بما فيها ملفات الإنجاز الشخصي التي تمثل عمل الطالب في شتى المواد والمقررات الدراسية.

الاختبارات التحصيلية الشفهية

تعد الاختبارات التحصيلية الشفهية التي يطلب فيها من الممتحن إعطاء إجابة شفهية على عدد من الأسئلة تطرح عليه شفهياً أو كتابياً من أقدم أنواع اختبارات تقييم الأداء. وإن أول تقرير لاختبارات كتابية في جامعة أوروبية كان من جامعة بولونيا عام ١٢١٩، ومع ذلك فقد استمر استخدام الاختبارات الشفهية لمئات السنين فيما بعد. وقد قام الرائد في مجال التعليم في الولايات المتحدة هوراس مان بإقناع لجنة مدينة بوسطون باستبدال اختباراتها الشفهية باختبارات كتابية في عام ١٨٤٥؛ إذ رأى مان أنه، وبالمقارنة مع الاختبارات الكتابية، لا تعد الاختبارات الشفهية عادلة في حق جميع المتقدمين للامتحان كما أنها تعاني من كونها اختبارات أقل شمولاً وتستغرق زمناً أطول ولا يتم توثيقها بالشكل اللازم وتشجع على المحسوبية وتفضيل بعض الممتحنين على البعض الآخر، هذا بالإضافة إلى صعوبة تقييم الأسئلة الصعبة أو الشديدة التعقيد. وعلى الرغم من اعتماد المؤسسات التعليمية في الولايات المتحدة على الاختبارات الشفهية في تقييم عدد من المهارات، تعد الاختبارات الكتابية كالاختبارات المقالية والاختبارات الموضوعية ذات الأسئلة متعددة الاختيارات وسائل التقييم الأساسية في المدارس.

وتعد طريقة الاختبار الشفهية أكثر شيوعاً في الجامعات الأوروبية منها في جامعات الولايات المتحدة الأمريكية. إلا أنه ومما لا شك فيه أن هذا النوع من الاختبارات فقد كثيراً من شعبيته في كلا القارتين. ولا تزال الاختبارات الشفهية مستخدمة في تقييم الأطفال في السنوات المدرسية الأولى حيث لا يزال الأطفال دون سن تسمح لهم بالقراءة والكتابة بسهولة. كذلك فإن هذه الاختبارات مبررة في الصفوف الأعلى وذلك عندما يصبح وقت التحضير والمرافق شحيحة. كما تستخدم الاختبارات الشفهية في مراحل الدراسة الإبتدائية المتقدمة وكذلك في المرحلة الثانوية إلا أنها تقتصر عادةً على مواد البلاغة والخطابة والمسرح

واللغات الأجنبية أو أكثر منها في المواد الأخرى. أما في المراحل العليا للدراسية الجامعية فتستخدم الاختبارات الشفهية كواحدة من متطلبات القبول في هذه المرحلة الدراسية أو في مناقشة الأطروحات التي يكتبها طلاب الماجستير والدكتوراة. كما تستخدم الأسئلة الشفهية في الفصول الدراسية في شتى المراحل التعليمية كجزء من العملية التعليمية إذ تتيح هذه الأسئلة الفرصة للطلاب في المشاركة الفعالة في الدرس وفي التعبير عن آرائهم والتواصل مع المدرس. كما يستخدم المدرسون والإداريون في المؤسسات التعليمية الأسئلة الشفهية لتقييم الطلاب وتحديد مدى تطورهم الأكاديمي مما يخدم عملية توزيع هؤلاء الطلاب على الصفوف الدراسية الملائمة أو عملية إعداد برنامج تعليمي خاص يناسب الحالات الخاصة لهؤلاء الطلاب.

كما تشكل الأسئلة الشفهية جزءاً من أسئلة مقابلات التوظيف ذات البنية الواضحة والعفوية إما وجهاً لوجه أو عبر الهاتف، وجزءاً من اختبارات التراخيص والشهادات المهنية والحرفية. ومما يساهم في رفع ثبات التقييم في هذه الحالة، توجيه الأسئلة ذاتها لكافة المتقدمين وتحديد المدة الزمنية التي يتوجب عليهم إتمام الإجابة خلالها وتسجيل الإجابات لمقارنتها لاحقاً مع إجابات باقي المتقدمين وتقييمها بالشكل المناسب.

يكره عدد كبير من الناس الاختبارات الشفهية ويعتبرونها اختبارات غير عادلة. إلا أن مدرسي الآداب واللغات الأجنبية والمسرح والخطابة يعزون الصعوبات التي يوجهها معظم الناس في التعبير عن أنفسهم شفهياً بالشكل الصحيح إلى قلة التركيز على المهارات الشفهية في كل من التدريس والاختبار (كراول وماكجينايتي، ١٩٧٤؛ بلات، ١٩٦١). ويقر معظم التربويين العاملين في مجال القياس النفسي بكون الاختبارات الشفهية أدوات متحيزة وقليلة الثبات والفعالية إذ غالباً ما تتطلب هذه الاختبارات وقتاً طويلاً ولا تغطي جزءاً مقبولاً من المقرر وتكون، في معظم الأحيان، ركيكة ووضعت على عجل.

وعلى الرغم من الانتقادات الشديدة للاختبارات الشفهية، تمتلك هذه الاختبارات ميزة توفيرها لفرصة التواصل المباشر بين الممتحَن ومعد الاختبار مما يسهل عملية تقييم مظهر الممتحَن وأسلوبه في الكلام. ومن الميزات الأخرى للاختبارات الشفهية: (١) صعوبة الغش والخداع فيها و(٢) توفر للقدرات الإدراكية الأكثر سمواً من تلك التي تتطلبها الاختبارات الكتابية و(٣) كونها فرص لتهيئة الممتحَنين للمواقف الاجتماعية التي تتطلب مهارة في التواصل الشفهي مع الآخرين و(٤) تشجع على مراجعة الطالب لمقرر الاختبار بحذر كبير و(٥) تتطلب الإجابة على أسئلتها وقتاً أقل من كل طالب و(٦) إن عمليتي وضع وتصحيح هذه الاختبارات لا تستغرقان وقتاً طويلاً كما هو حال الاختبارات الموضوعية. كما يمكن لواضع الاختبار أن يطلب من الممتحن أن يوضح بعض النقاط أو يتعمق فيها أو أن يسأل أسئلة أخرى تهدف إلى تحدي الممتحن أو سبر قدرته على التحليل والتفكير الإبداعي.

وتدل نسبة الارتباط المتوسطة بين درجات الأفراد في الاختبارات الشفهية والاختبارات الكتابية للمعلومات ذاتها على أن الاختبارات الشفهية تقيس نواحٍ أخرى من تحصيل الفرد في مجالٍ ما. وكما كان الحال مع الاختبارات التحصيلية القياسية واختبارات التحصيل التي يضعها المدرسون والتي تحدثنا عنها في الفصل الرابع، تعطي اختبارات التحصيل الشفهية معلومات تختلف عن اختبارات التحصيل الكتابية. وبشكلٍ عام تعد اختبارات التحصيل الموضوعية أقدر على قياس معرفة الطالب بحقائق ومعلومات محددة، بينما تساهم الاختبارات الشفهية والاختبارات المقالية في تقييم عمق هذه المعرفة ومداها.

وكما هو الحال في عملية المقابلات فإن أهم مصادر عدم الثبات أو الأخطاء التي تعترض نجاح الاختبارات الشفهية هي الأفراد الذين يسألون الأسئلة أنفسهم. إذ يشترط في هؤلاء الأفراد أن يكونوا على قدر كبير من المعرفة والخبرة في مادة

الاختبار وأن يكونوا على معرفة تامة بهدف السؤال ومتطلبات الإجابة. كما يجب عليهم أن يصغوا بانتباه وحذر شديدين وأن يملكوا القدرة على استدراج المتحَنين للبوح بكل ما يعرفونه عن موضوع الاختبار. كما يجب على هؤلاء الأشخاص عند تقييم إجابات المتحَنين مراعاة ضرورة التعرف على أنماط محددة للسلوك التي لاحظوها وعدم الاكتفاء باختيار عبارات منمقة ومبهمة في الوقت نفسه للتعبير عن أداء المتحَن. فعلى سبيل المثال، استخدام عبارات مثل: "عنده قدرات إبداعية كامنة" أو "يتمتع بقدرة عقلية جيدة" أو "قادر على التواصل بشكل فعال" لا يقدم ولا يؤخر في تقييم أداء الطالب بشكل دقيق. وتتيح الاختبارات الشفهية فرصة تقييم المتحن ليس من حيث تنظيم إجابته والمعلومات التي تضمنتها الإجابة بل من حيث مزايا الطالب الشخصية وأسلوبه وطريقة كلامه (إرتفاع أو إنخفاض الصوت، نبرة ونغمة الصوت، سلامة اللفظ والنطق، واستخدام المفردات والعبارات). وتتوفر عدد من إستمارات التقييم التي يمكن أن تساعد في زيادة ثبات وموضوعية الاختبارات الشفهية (غرين، ١٩٧٥؛ غويرا وأبرامسون ونيومارك، ١٩٦٤؛ وراجع الشكل ٥-١).

تفيد التوصيات التي تحدثنا عنها في الفصل الثالث فيما يخص عملية وضع درجات الأسئلة المقالية في تقييم إجابات الأسئلة الشفهية. فيجب أولاً وضع نموذج عن الإجابة المثالية للسؤال (وذلك قبل بدء الاختبار) ثم الإستعانة بعدد من المحترفين المدربين لوضع الدرجات (مما يزيل خطر المحسوبية وتفضيل ممتحن على آخر). كما يمكن تعزيز ثبات الاختبارات الشفهية وصدقها بتسجيل الإجابات والعودة إليها لاحقاً لإجراء إعادة تقييم لعملية التقييم الأولية. وتتوفر عدد من الإجراءات الإحصائية التي يمكن أن تفيد في تصحيح أخطاء التقييم لإجابات الأسئلة الشفهية (راجع رايموند وويب وهيوستون، ١٩٩١).

ويختلف اختبار مهارات الإستماع والحديث عن الاختبارات الشفهية التقليدية كون الأولى تركز على مهارات الإستماع والتحدث وليس على فهم

ومناقشة الأفكار المختلفة. ويمكن للقارئ الحصول على قائمة بعدد من الاختبارات القياسية التي تقيس قدرات الأطفال والبالغين اللغوية الشفهية باختيار القائمة ٣ من برنامج F-8 الموجود على القرص المرافق لهذا الكتاب. وتُستخدم اختبارات الإستماع والحديث بشكلٍ خاص لقياس القدرة على تعلم اللغات الأجنبية أو لغة ثانية. ومن الأمثلة على هذه الاختبارات مؤشر التقييم اللغوي (الشفهي) الذي يقيس مهارات الإستماع والحديث في اللغتين الإنجليزية والإسبانية والذي يساعد في إلحاق الطلاب بصفوف اللغة الإنجليزية كلغة ثانية أو إستبعادهم منها.

الشكل ٥-١

قائمة بما يجب فعله عند تقييم تقرير شفهي

تعليمات عامة: ضع إشارة X إلى جانب الجمل التي تعبر عن تصرفات ومظهر الطالب خلال إعداد التقرير الشفهي.

أولاً: المظهر العام للطالب:

١- كان الطالب منحني القامة.

٢- كان الطالب منتصب القامة.

٣- كانت حركات الطالب مرتبكة وغير متناسقة.

٤- كانت حركات الطالب متناسقة.

٥- كانت يدا الطالب ترتجفان.

٦- كانت يدا الطالب ثابتتين.

٧- كانت ملامح وجه الطالب تعبر عن ملله أو كانت خالية من التعبير.

٨- كانت ملامح وجه الطالب مفعمة بالحيوية وتدل على اهتمامه بما يتحدث عنه.

ثانياً: ملاحظات عامة على الطالب

٩- بدا الطالب مرتبكاً.

١٠- بدا الطالب مرتاحاً ومسيطراً على الموقف.

١١- لم يتواصل الطالب بالنظر إلا مع عدد قليل من الأفراد.

١٢- حقق الطالب التواصل بالنظر مع كل جمهوره.

١٣- قرأ الطالب ملاحظته كما هي ولم يعد صياغتها.

١٤- تحدث الطالب بصوت متردد وغير متناغم.

١٥- تحدث الطالب بصوت هادئ ومعبر.

١٦- كان صوت الطالب مرتفعاً.

١٧- تحدث الطالب بصوت خافت.

١٨- تحدث الطالب ببطء شديد.

١٩- تحدث الطالب بسرعة شديدة.

٢٠- كانت مخارج الحروف عند الطالب سيئة.

٢١- توقف الطالب طويلاً مراراً أثناء الحديث.

٢٢- استخدم الطالب مصطلحات عامية.

٢٣- كرر الطالب عدداً من الأفكار بلا ضرورة.

٢٤- عبر الطالب عن جميع أفكاره.

١٥- قدم الطالب المعلومات بشكل منظم.

٢٦- أثبت الطالب فهمه لموضوع التقرير.

٢٧- لخص الطالب أفكاره بشكل جيد.

٢٨- أجاب الطالب على أسئلة الحضور بشكل جيد.

التقييم العام لأداء الطالب على مقياس من ١ إلى ١٠ يمثل العدد فيه ١ أفضل أداء والعدد ١٠ أسوأ أداء:.........................

الملاحظات:

....

....

....

....

....

اختبارات الأداء والبورتفوليو (الملفات)

إن معرفة وفهم القوانين والأساليب لعمل مهاري لا تعد كافية للأداء الفعال إذ يجب ترجمة ما تم إكتسابه إدراكياً إلى عدد من الأفعال يقوم بها شخص يمتلك الدافع الكافي لجعل مستوى أدائه مقبولاً. وعليه فإن نجاح طالبٍ ما في أي اختبار موضوعي له علاقة بمواد مدرسية أو حرفة معينة، لا يضمن نجاح هذا الطالب في تحويل معرفته النظرية إلى أداة فعالة تقوده إلى النجاح في العالم المعقد الذي نعيش فيه، ولذلك تم العمل على البحث عن طرق أخرى لتقييم الأداء تتسم بالواقعية والعمومية وتتألف من مهام تشبه إلى حد كبير السلوك البشري المراد تقييمه وتسمح للمجيب على السؤال بالكشف عن قدراته ومعرفته.

وإذا قبلنا تعريف "الأداء" على أنه أي سلوك أو تصرف يقوم به الفرد فإن الاختبارات ذات الأسئلة متعددة الاختيارات هي نوع من اختبارات الأداء التي تتطلب من المتحَن القيام بمهمة ما حتى وإن كانت هذه المهمة هي اختيار الإجابة الصحيحة. إلا أنه يمكن القول كذلك أن اختبارات الاختيار المتعدد وغيرها من الاختبارات الموضوعية تتصف بالسطحية الشديدة من حيث ما تطلب من الطالب

القيام به (اختيار الإجابة الصحيحة) ولا تختبر الحدود الحقيقية لإمكانيات الفرد ومعرفته. ولهذا تعد الاختبارات المقالية والاختبارات الشفهية أدوات قياس أكثر عمقاً من الاختبارات الموضوعية إذ أن إجابات الطلاب المقالية والشفهية تكون عادة إجابات أطول وأكثر قدرة على التحليل والتعمق. كما تفيد الاختبارات المقالية والشفهية في تقييم قدرة الطالب على تنظيم أفكاره وصياغتها بطريقة واضحة وسليمة.

ويطلق اسم "التقييم الواقعي" Authentic أو "التقييم البديل" على اختبارات الأداء وذلك للتأكيد على الواقعية وعلى إختلاف هذه الاختبارات عن الاختبارات الموضوعية التقليدية. ومن الجدير بالذكر أن اختبارات الأداء لا تقتصر على مجالات التعليم الزراعي والفنون المسرحية والتربية البدنية وتدريب السائقين وعمال المحال التجارية بل تتعداها إلى مجالات المواد الأكاديمية كاللغة الإنجليزية والرياضيات والعلوم الطبيعية والعلوم الاجتماعية وغيرها. ويعود سبب استخدام اختبارات الأداء في هذه المواد إلى قناعة مدرسيها بضرورة تحويل المعلومات التي يتلقاها الطالب داخل الصف إلى مهارات تفيده في مجال العمل والحياة العملية. وتهيئ المدارس لطلابها عادة فرصة إكتساب عدد من المهارات التي تساعدهم في الحياة العملية كالقراءة والكتابة والتحدث والرسم وممارسة الرياضات المختلفة والغناء في فرقة المدرسة للإنشاد والعزف على الآلات الموسيقية وتصميم الأشياء باستخدام الأدوات والآلات والطباعة وغيرها من المهارات المكتبية (انظر الشكل ٥-٢ وبرنامج F-7 على القرص المرفق بالكتاب). كما تعد المشاريع التي يعمل عليها الطلاب والتطبيقات العملية التي يقومون بها في مختبرات العلوم والتدرب على التحدث أمام الجمهور باللغة الإنجليزية واللغات الأجنبية الأخرى وحل المشاكل في مجالات العلوم الطبيعية والإنسانية والآداب والرياضيات نشاطات مفيدة في الحياة العملية. ويوضح الشكل ٥-٣ مثالاً من اختبار أداء في مجال العلوم. ويتطلب هذا النوع من المهام الرياضية أو العلمية من الطالب صياغة نص

المسألة وتنظيم أفكاره وجمع أنواع مختلفة من الأدلة بالإضافة إلى القدرة على الإبداع والإبتكار. أي أن هذا النوع من الأسئلة يقيِّم الطالب بطريقة واقعية لا تقدر عليها اختبارات الورقة وقلم الرصاص المعتادة. ولا يعد استخدام هذا النوع من الاختبارات أمراً جديداً إذ لطالما استخدمت المؤسسات التعليمية والشركات اختبارات مشابهة، إلا أن ما يعتبر جديداً هو استخدام هذا النوع من الاختبارات على مستوى الولايات كجزء من برامج للتقييم تهدف إلى اختبار الطلاب باستخدام أسئلة ومهام تعليمية واقعية وذات فائدة عملية.

الشكل ٥-٢

عينة من أسئلة اختبار SRA5 للطباعة. (أعيدت طباعته بموافقة دار نشر ماكغرو - هيل / لندن)

جزء من ورقة الاختبار (أ) - الطباعة السريعة

عزيزي السيد كولمان:

أرفق طياً عدد من النسخ لمقترح شراء أرض مساحتها خمس فدانات تم اختيارها لبناء مصنعنا الجديد عليها. أرجو أن توزع النسخ على جميع العاملين لديك.

جزء من ورقة الاختبار (ب) - المراسلات التجارية

السيدة ل. آشمان

٢٨٢٤ ليروي درايف

ألبرتا، إلينوي ٠٢١٠١

عزيزتي السيدة آشمان:

لقد كان اجتماعي مع ممثلنا في المنطقة الوسطى اجتماعاً ناجحاً إلا أنه لا تزال توجد عدد من النقاط التي يجب علينا ...

جزء من ورقة الاختبار (ج) – المهارات العددية

عزيزي السيد بينك:

يوضح الجدول أدناه حركة البيع لدينا في الأشهر الأخيرة. وكما تلاحظ لقد تأرجحنا كثيراً في هذه الفترة إلا أنني على ثقة من أن الوضع سوف يستقر في الربع القادم من السنة المالية.

	يناير	فبراير	مارس	المجموع
الصمامات	١٧,٥	١٢,٢	١٥,٣	٤٥,٠
القواطع	٩,٣	٦,١	٢,١	١٧,٥

تصنيفات السلوك الحركي – النفسي

ينطوي التحضير لوضع اختبار يقيس مهارة ما على وضع قائمة مفصلة بالتصرفات والأفعال التي تؤشرعلى درجات كفاءة متفاوتة في هذه المهارة. ويمكن الإستعانة بعدد من المراجع عند إعداد اختبارات في مجالات التربية البدنية والرياضات المختلفة وعلم التمارين الرياضية (مثل سافريت ووود، ١٩٩٥؛ ستراند وويلسون، ١٩٩٣).

الشكل ٥-٣

مثال توضيحي على مهمة لتقييم الأداء في مادة العلوم (من كتاب عينة، ١٩٩٢)

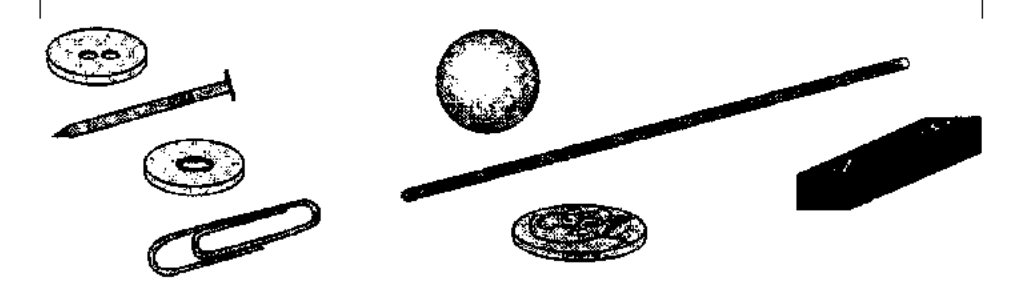
المغناطيس

وصف المهمة:

استخدام المغناطيس للتعرف على الأشياء التي يمكن جذبها وتلك التي لا يمكن جذبها ثم تفسير الفرق بينهما.

الأدوات المواد المستخدمة:

معناطيس، والعناصر السبعة التالية: زر بلاستيكي، حلقة مصنوعة من الفولاذ أو الحديد، مشبك للورق، مسمار حديدي، كرة زجاجية صغيرة، عصا بلاستيكية، قطعة نقود نحاسية.



تعليمات للطالب:

اختبر الأشياء من حيث إنجذابها نحو المغناطيس أو عدم إنجذابها ثم قسم الأشياء إلى مجموعتين وفسر الإختلاف بين العناصر في كلا المجموعتين.

طريقة وضع الدرجات:

تعطى درجات للطلاب الذين قسموا الأشياء بالشكل الصحيح. لقد تم تسجيل أربعة شروح وهي: أن مجموعة كانت مصنوعة من الحديد أو الفولاذ، وأن مجموعة أخرى انجذبت نحو المغناطيس، وأن مجموعة ثالثة كانت مصنوعة من الحديد والفولاذ هي التي جذبها المغناطيس، والملاحظات الأخرى تشكل المجموعة الرابعة.

كما تعد تصنيفات الأهداف في المجال النفسي – الحركي مفيدة كذلك (هارو، ١٩٧٢؛ نيكسون وجويت، ١٩٨٠؛ سيمبسون، ١٩٦٦). ويشابه الإطار العام لتصنيف نيكسون وجويت للسلوك الحركي كتاب تصنيف الأهداف التعليمية: المجال الإدراكي (بلوم وكراثوول، ١٩٥٦) و كتاب تصنيف الأهداف التعليمية:

المجال العاطفي (بلوم وكراثوول وماسيا، ١٩٦٤). وتعد التصنيفات التي وضعها نيكسون وجويت هرمية ويحتوي الأكبر منها على الأصغر بالتتابع وتهتم بالعملية وليس بالمحتوى. فعلى سبيل المثال، تعنى الطبقات الدنيا من التصنيف بالمفاهيم التي يكونها الفرد عن العناصر التي تؤلف حركة معينة وبالمحاولات التي يقوم بها هذا الفرد لإيجاد نمط لهذه الحركة على أساس ملاحظته أو تذكره لها. وتعنى الطبقات العليا من التصنيف بالحركات التي يبتكرها الأفراد في مواقف معينة يتعرضون لها أثناء ممارسة الرياضة أو الرقص أو غيرها من النشاطات الحركية. ويتطلب تطبيق تصنيف الأهداف التعليمية للاختبار أو وضع قائمة مفصلة بهذه الأهداف في المجال النفسي - الحركي تحديد الثقل المعطى لكل عنصر من عناصر الأداء وتحديد ما إذا كان بطء الطالب أو قلة ترتيبه أو الأخطاء التي يرتكبها سوف تؤدي إلى حذف عدد من الدرجات التي تم منحه إياها. ومما لا شك فيه أن للاهتمام بالعناصر المنفصلة التي تكوّن الأداء فوائد تشخيصية وعلاجية وتقييمية في العملية التعليمية.

كما يجب الربط بين الأهداف التعليمية أو المحكات للأداء في المجال الإدراكي وبين تصنيفات الأفعال وأنماط السلوك أو ما يعرف بنظرية التعلم. ويجب التعرف على هذه الأهداف أو المحكات كجزء من مجهود مشترك يبذله المدرسون والطلاب في الصفوف الدراسية قبل البدء بالعملية التعليمية والوصول إلى النتائج المرجوة منها. كما يمكن أن يشارك الطلاب في وضع اختبار الأداء واختيار الأسئلة والمهام التي ستدخل في الاختبار. وتشجع هذه الممارسة الطلاب على أخذ دور فاعل في العملية التعليمية كما أنها تحول عملية الاختبار إلى عملية ذات معنى بالنسبة للطالب.

تقدير الأداء

تفيد تسجيلات الفيديو في إثبات قدرة فرد ما على ممارسة مهارة ما بشكل جيد، كما أنها تساعد على مراقبة تقدم المتعلِّم وتقييم هذا التقدم. ويتطلب تقييم

الأداء في المجالين النفسي – الحركي والإدراكي النظر إلى المنتج النهائي بالإضافة إلى العملية أو الخطوات التي قادت إليه. إذ لا يكفي تقييمنا لأداء لاعب الغولف مثلاً بناءً على عدد الكرات التي نزلت في الحفر أو للاعب كرة السلة بناءً على عدد الرميات الحرة والنقاط التي كسبها أو للاعب التنس بناءً على عدد الأشواط التي ربحها، إذ يجب أخذ أسلوب وأسلوب اللاعب في اللعب وحتى طريقة تعامله مع الآخرين بعين الإعتبار. وعليه، لا يتم تقييم المنتج النهائي الذي يقدمه الطالب فحسب بل يجب أخذ عوامل أخرى ككمية ونوعية المنتج ودرجة إتقان عملية الإنتاج في عين الإعتبار.

وغالباً ما تتم عملية تقييم المنتج والعملية الإنتاجية في المجالين النفسي – الحركي والإدراكي بشكلٍ ذاتي عن طريق الرجوع إلى عدد من التعليقات والروايات (عن عمل الطالب) وقوائم الرصد وسلالم التقدير. ويجب إعداد سلم الدرجات الذي يسهل عمل المقيِّم قبل إجراء الاختبار. ثم يراقب المدرس أو المسؤول عن التقييم أداء الطالب مباشرةً أو عن طريق شريط الفيديو ويسجل ملاحظاته وتقديره للأداء. ويمكن الرجوع إلى الكتابين الأول والثاني من هذه السلسلة (أيكين، ١٩٩٦) للحصول على معلومات تفصيلية عن إعداد وقوائم الرصد وسلالم التقدير المستخدمة في تقييم الأداء. ومن المهم جداً أن تسجل الملاحظات وأن يتم تقدير الأداء بحذر ودقة شديدين مع مراعاة مصادر التحيز التي يمكن أن تتدخل في عملية الحكم على أداء الطالب. ويتطلب هذا إعداد سلم درجات محكم يحدد متطلبات الأداء على مختلف المستويات، كم يتطلب أن تكون المشاريع التي يعمل عليها الطلاب رفيعة المستوى. ويجب على من يقوم بعملية التقييم أن يكون مدرباً أو حاملاً للمؤهلات المطلوبة والتي تم التحقق منها (ستيغينز، ١٩٩٤). ويمكن لعملية وضع الدرجات أن تكون عملية شمولية تنظر إلى أداء الطالب بشكلٍ عام أو تحليلية تقسم الأداء إلى عناصر مختلفة يتم تقييم كل عنصر منها على حدة.

مزايا اختبارات الأداء وعيوبها

تمتلك اختبارات الأداء الإدراكية عدداً من المزايا والعيوب تختلف عن اختبارات الورقة وقلم الرصاص التقليدية.

ويعد عدم اكتفاء اختبارات الأداء بالإشارة العابرة من قبل الممتحَنين إلى حدود معرفتهم وتركيز هذه الاختبارات على قيام الطالب بتأدية نمط سلوكي معقد يستمر لفترة طويلة من الزمن أحد أهم مزايا اختبارات الأداء. كما يعد تركيز اختبارات الأداء على المهارات المعقدة وعلى ضرورة اقتران إجابات الطلاب بتفسيرات وتبريرات لإجاباتهم عاملاً مساعداً على تقييم الطالب والعملية التربوية في آنٍ واحد يمكن أن يقود إلى تناغمٍ تام بين عمليتي التدريس والاختبار. ولهذا السبب تبدو اختبارات الأداء المحكمة الصنع وكأنها مهمات درسية. ولعل أحد أهم المزايا التي تتمتع بها اختبارات الأداء كونها عملية وواقعية ووقتها مفتوح وتتطلب تعاوناً بين المدرسين والطلاب وبين مجموعات مختلفة من الطلاب.

ويمكن للمدرس أن يعد تقريراً واحداً عن أداء مجموعة من الطلاب باستخدام قائمة رصد جماعية. كما يمكن إعداد قائمة رصد أو سلم تقدير لأداء كل طالب يتم فيه تقييم مشاركة الطالب في الحوار مع المجموعة ومساهمته بالمعلومات وإصغائه إلى ما يقوله الآخرون وتقيده بالتعليمات وتقبله لآراء الآخرين وقبوله ثم تنفيذه للمهام الموكلة إليه من قبل المجموعة (كوبيسزين وبوريخ، ١٩٩٠).

ومما لا شك فيه أن كون اختبارات الأداء اختبارات مفتوحة (لا يمكن حصر إجاباتها أو تحديدها) يجعل مهمة وضع درجاتها أكثر صعوبة، إلا أنه وفي الوقت ذاته تعتبر خاصية المرونة هذه التي تتمتع بها اختبارات الأداء أكثر قدرة على تقييم الأفراد في مواقف حقيقية مستمدة من الواقع العملي حيث يُتوقع من

الأفراد الجمع بين كمٍ كبير من المعلومات قبل التوصل إلى حلول للمشكلات أو قبل إبداع شيء جديد. وتقود ضرورة الحصول على معلومات من مصادر مختلفة إلى إثبات الفرد قدرته على تنظيم المعلومات والجمع بينها وتقييمها قبل إتخاذ القرارات. إلا أن هذا يستغرق وقتاً طويلاً وقد تكون خطواته غير واضحة دائماً.

وتعد اختبارات الأداء ذات البنية المنظمة والتي يتم اختبار جميع الممتحَنين فيها ضمن الظروف ذاتها أكثر موضوعية من الاختبارات العفوية (أي التي لا تتميز ببنية واضحة) التي تقيم عدداً من الطلاب بشكل جماعي أثناء تأديتهم لمهام معينة في جو من الحرية المطلقة. ومن الممكن أن يقوم عدد من الأفراد بتقييم أداء الطلاب أو أن يقوم الفرد ذاته بتقييم أداء الطالب عدة مرات ويتم بعد ذلك حساب متوسط درجات جميع عمليات التقييم. مع هذا كله وعلى الرغم من محاولة تقييم اختبارات الأداء بشكل عادل وبعيد عن التحيز، تبقى درجات اختبارات الأداء أقل ثباتاً من الاختبارات الموضوعية.

يتم التحقق من ثبات اختبارات الأداء بشكل نموذجي عن طريق الطلب من شخصين أو أكثر تقييم المُنتج الذي توصلت إليه مجموعة من الطلاب أو تقييم العملية التي توصلوا من خلالها إلى المنتج، ثم يتم حساب معامل الإرتباط لمجموع المقيمين كمؤشر لثبات التقييم باستخدام برنامج E-4. كما يمكن كذلك الطلب من عدد من المقيمين تقييم أداء الطلاب أو منتجاتهم كل على حدة، ثم، وباستخدام الطريقة التي تحدث عنها واينر وبراون وميشلز (١٩٩١) أو باستخدام برنامج E-5 يتم حساب معامل المطابقة Coefficient of Concordance (بين الدرجات المختلفة التي وضعها المقيمون المختلفون).

تحتاج اختبارات الأداء مقارنة باختبارات الورقة وقلم الرصاص إلى وقت طويل وإلى معدات باهظة الثمن وإجراءات معقدة، هذا بالإضافة إلى تدني مستوى ثباتها (والذي تحدثنا عنه سابقاً). وعليه، يمكن الإستعاضة عن اختبارات

الأداء باختبارات مقالية إن لم تكن الواقعية التي تتميز بها اختبارات الأداء هي الاعتبارالأهم لدى واضعي الاختبار. كما أن اختبارات الأداء لا تعالج مشكلة التحيز التي تعاني منها اختبارات الورقة وقلم الرصاص (أو الاختبارات ذات أسئلة الاختيارات المتعددة)، إذ لا يمكن القول أنه لا أثر للتحيز إلى جنسٍ أو عرقٍ ما في اختبارات الأداء. وعلى الرغم من أن اختبارات الأداء لا تتأثر بما تحدثنا عنه سابقاً من براعة بعض الأفراد في مجال الاختبارات، يمكن لعوامل أخرى كالمهارات اللغوية والمعرفة السابقة للفرد وخبراته أن تكون مصدراً لتحيز اختبارات الأداء لبعض المجموعات البشرية على حساب مجموعات أخرى. وبالرغم من الحماس الشديد الذي يمتلكه البعض لاختبارات الأداء، لا يمكن القول بأن هذه الاختبارات هي الإجراء الأمثل في عالم الاختبارات. وفي الممارسات الحالية يتم تقديم الحلول الوسطية التي تجمع بين اختبارات التحصيل القياسية التقليدية وأدوات التقييم التي تستند إلى أسس تقييم الأداء.

اختبارات الأداء المتوفرة في الأسواق

توزع الشركات التالية: ماكغرو – هيل (CTB) وشركة ريفرسايد للنشر ومؤسسة علم النفس المتحدة عدداً من بطاريات اختبار الأداء وملاحق لبطاريات اختبارات التحصيل القياسية. ويضم الجدول ٥-١ قائمة بهذه الاختبارات ووصف بسيط لها وأسماء ناشريها والسنوات الدراسية التي يمكن استخدامها معها.

الجدول ٥-١

قائمة بعدد من اختبارات الأداء المتوفرة في الأسواق

- اختبار AMAP لتقييم الأداء في مجال المهارات اللغوية للصفوف ١ إلى ٣، الناشر: مؤسسة علم النفس المتحدة. نبذة: يتألف الاختبار من مهام مستوحاة من الحياة العملية تتطلب قراءة النصوص الأدبية والتفكير بمعانيها ثم الإجابة على أسئلة متعلقة بها بشكل فردي أو ضمن مجموعات صغيرة.

- اختبار AMAP لتقييم الأداء في مجال البحث في الرياضيات للصفوف ٣ إلى ١٠، الناشر: مؤسسة علم النفس المتحدة. نبذة: يتألف الاختبار من مهام مستوحاة من الحياة العملية تدفع الطلاب إلى القيام بتحقيقات في مجال الرياضيات تقودهم إلى فهم وتقدير قيمة الرياضيات وفائدتها في الحياة العملية.
- ملحق الإجابات الإنشائية لاختبارات أيوا للصفوف ٣ إلى ١١ أو ١٢، الناشر: شركة ريفرسايد للنشر. نبذة: يقييم الاختبار أداء الطلاب في مجالات القراءة واللغة والرياضيات.
- اختبار CTB لتقييم الأداء للصفوف ٢ إلى ١١، الناشر: ماكغرو – هيل (CTB). نبذة: يتألف الاختبار من تمارين في القراءة والمهارات اللغوية والرياضيات والعلوم الطبيعية والاجتماعية.
- اختبار GOALS: مقياس التحصيل الذي يعتمد على تقييم الأداء للصفوف ١.٥ إلى ١٢، الناشر: مؤسسة علم النفس المتحدة. نبذة: يركز الاختبارعلى عدد من النواحي التعليمية كالقراءة (النصوص الروائية والإعلامية) والرياضيات (حل المسائل والإجراءات الرياضية) واللغة (النصوص الأدبية والإيضاحية) والعلوم (علم الحياة وعلوم الأرض والفضاء) العلوم الاجتماعية (الأسس الجغرافية والاقتصادية والحضارية والتاريخية والسياسية للمجتمع البشري).
- اختبار IAS لتقييم الأداء في مجال المهارات اللغوية (تتوفر نسخة باللغة الإنجليزية وأخرى بالإسبانية) للصفوف ١ إلى ٨، الناشر: مؤسسة علم النفس المتحدة. نبذة: يتألف الاختبار من مقاطع نصية للقراءة وعدد من تمارين الكتابة الموجهة تهدف إلى تقييم فهم النصوص والقدرة على التعبير الكتابي بالإضافة إلى مهارات التفكير العليا.
- اختبار IAS لتقييم الأداء في مجال الرياضيات للصفوف من ١ إلى ٨، الناشر: مؤسسة علم النفس المتحدة. نبذة: يتألف الاختبار من عدد من المسائل والمواقف التي تتطلب حلولها استخدام عدد من المهارات الرياضية.

- اختبار IAS لتقييم الأداء في مجال العلوم للصفوف من ١ إلى ٨، الناشر: مؤسسة علم النفس المتحدة. نبذة: يتألف الاختبار من عدد من الصور المترافقة مع وصف بسيط للمشكلة أو المهمة التي يتعين على الطالب أن يبحث عن حل لها يتوصل إليه بالطريقة التي يراها مناسبة.
- اختبارات المستوى العام بالنسبة إلى المقاييس التعليمية الجديدة للمراحل الإبتدائية والمتوسطة والثانوية. الناشر: مؤسسة علم النفس المتحدة. نبذة: يقييم الاختبار مهارات اللغة الإنجليزية والرياضيات.
- اختبارات تقييم الأداء ITBS وTAP وITED للصفوف ١ إلى ١٢، الناشر: شركة ريفرسايد للنشر. نبذة: يقيِّم الاختبار مهارات اللغة الإنجليزية والرياضيات والعلوم الطبيعية والاجتماعية.
- سلسلة ريفرسايد لتقييم الأداء للصفوف ١ إلى ١٢، الناشر: شركة ريفرسايد للنشر. نبذة: يقيِّم الاختبار الإجابات الإنشائية في مجالات القراءة والكتابة والرياضيات.
- اختبار ستانفورد ٩ المفتوح في مجال الرياضيات للصفوف ١,٥ إلى ١٣، الناشر: مؤسسة علم النفس المتحدة. نبذة: يتألف الاختبار من تسع أسئلة أو مهام مرتبطة بموضوع واحد. ويتم وضع درجات الأسئلة بناءاً على مؤشر مؤلف من أربع نقاط يقيِّم المعايير التالية: طريقة التعبير وطريقة التفكير وحل المشكلات.
- اختبار ستانفورد ٩ المفتوح في مجال القراءة للصفوف ١,٥ إلى ١٣، الناشر: مؤسسة علم النفس المتحدة. نبذة: يتألف الاختبار من مقاطع نصية روائية يتبعها ٩ أسئلة مفتوحة.
- اختبار ستانفورد ٩ المفتوح في مجال العلوم للصفوف ١,٥ إلى ١٣، الناشر: مؤسسة علم النفس المتحدة. نبذة: يتألف الاختبار من تسع أسئلة أو مهام مرتبطة بعلوم الحياة وعلوم الأرض والعلوم الفيزيائية التي يتطلب إنجازها تطبيق عدد من المفاهيم وأتخاذ القرارات وحل المشكلات والبحث العلمي.

● اختبار ستانفورد ٩ المفتوح في مجال العلوم الاجتماعية للصفوف ١،٥ إلى ١٣، الناشر: مؤسسة علم النفس المتحدة. نبذة: يتألف الاختبار من تسع أسئلة واقعية صممت التقيّم فهم الطالب لعدد من الأمور التاريخية والجغرافية والإقتصادية والسياسية والحضارية.

وتعد هذه الاختبارات أدوات قياس مفتوحة (أي إجاباتها غير محدودة) تقدم للطالب عدداً من التمارين في التفكير الإستنتاجي وبخاصة في نواحي المهارات الأساسية بهدف اختبار قدرة الطالب على التفكير الإستراتيجي وعلى حل المشكلات بمفرده أو ضمن مجموعات صغيرة. وتتألف غالبية التمارين من مشكلات مستمدة من الحياة العملية يجب على الطالب، ومن خلال إجابات مطولة مبنية على المعلومات التي أعطيت له، إيجاد حلول لها أو محاولة معالجتها. وتتوفر من بين هذه الاختبارات اختبارات ذات مرجعية محكية وأخرى ذات مرجعية معيارية.وعلى حين تكتفي بعض هذه الاختبارات بتقييم القراءة والإمكانيات اللغوية والرياضيات والعلوم والعلوم الاجتماعية، تتوفر اختبارات (كاختبار CTB لتقييم الأداء واختبار GOAL واختبارات المستوى العام بالنسبة إلى المقاييس التعليمية الجديدة واختبارات تقييم الأداء ITBS وTAP وITED وسلسلة ريفرسايد لتقييم الأداء) تقيم الأداء في مادتين أو أكثر في آنٍ واحد.

وتعد الاختبارات الثلاثة التالية أكثر اختبارات الأداء شمولاً: اختبار GOALS: وهو مقياس التحصيل الذي يعتمد على تقييم الأداء، واختبارات تقييم الأداء ITBS وTAP وITED، واختبار CTB لتقييم الأداء. وتُقيِّم هذه الاختبارات أداء الطلاب في مهارات اللغة الإنجليزية والرياضيات والعلوم الطبيعية والاجتماعية حيث يطلب من طلاب الصفوف ١ إلى ١٢ أن يكتبوا إجابات أو يرسموا مخططات أو يضعوا عناوين للصور أو يكملوا الجداول أو ينقحوا النصوص، وذلك في مواد دراسية ترتكز على المعلومات أو على الخطوات

العملية[١] (كالقراءة والرياضيات واللغة والعلوم والعلوم الاجتماعية). وتمكن هذه الاختبارات الطلاب من إثبات قدرتهم على تطوير مهارتهم بشكلٍ يسمح لهم أن يكونوا طلاب علم جادين وفعالين. وتعكس معظم اختبارات GOAL التغيرات التي طرأت على المناهج التعليمية وعلى طرق التدريس كحركة اللغة الشاملة في مجال تعليم اللغات ومعايير NCTM للتحصيل العلمي في مجال الرياضيات وأهداف مشروع ٢٠٦١ في تدريس مادة العلوم. ويجمع الطلاب في اختبارات CTB لتقييم الأداء بين معلومات من مصادر متعددة أو من نص واحد مأخوذ من مصدر واقعي (كالجرائد والمجلات وغيرها) للتوصل إلى استنتاجات أوتخمين إجابات أو تحديد أفضل الطرق للإجابة على السؤال أو حل المشكلة، وذلك في مجالات القراءة والمهارات اللغوية والرياضيات والعلوم والعلوم الاجتماعية. ومن الجدير بالذكر أن التمارين التي تتضمنها سلسلة اختبارات ريفرسايدللأداء (R-PAS) في مجالات القراءة والكتابة وتمارين الرياضيات تتميز بالأصالة وبكونها تحفز على التفكير؛ وتتوفر اختبارات هذه السلسلة باللغتين الإنجليزية والإسبانية.

البورتفوليو (الملفات) Portfolios

لقد تم أخذ كلمة portfolio والتي تعني حرفياً "الملف أو المحفظة التي تحوي أوراق" من عالم السياسة والأعمال لتستخدم في مجال التعليم والتقييم التعليمي، حيث تشير هذه الكلمة إلى مجموعة من المشاريع والمنتجات والعروض والرسومات وغيرها من الأعمال التي قام بها الطالب خلال فترة زمنية معينة والتي يمكن على أساسها تقييم التقدم الأكاديمي الذي أحرزه الطالب. ويتألف

(١) يمكن تقسيم المواد التي يدرسها الطالب إلى مواد يعتمد التقييم فيها على سبر معلومات الطالب ومواد يعتمد التقييم فيها على متابعة عمل الطالب والخطوات التي يتبعها في إنجاز مشروع ما. فعلى سبيل المثال يُقيم الطلاب في مادة الرياضيات عن طريق سبر المعلومات أو مراقبة الخطوات التي يتخذونها لإيجاد الحلول. (ملاحظة المترجم).

بورتفوليو الكتابة من عينات من كتابات الطالب، وبورتفوليو الرياضيات من عدد من المسائل أو المشاريع أو الأحاجي التي أتمها الطالب أو أوجد حلولها (مجلس تعليم ولاية فيرمونت، ١٩٩١)، وبورتفوليو العلوم من مشاريع وتجارب وتقارير معملية وغيرها من المهام التي حققها الطالب. ولا تقتصر البورتفوليوعلى مادة واحدة إذ يمكن أن تضم نماذج من أعمال ومشاريع الطالب في عدد من المواد المدرسية أو المواد المرتبطة بالدراسة. فمن الجائز أن يضم البورتفوليوعينة من كتابات الطالب (كالمقالات أوالقصائد أوغيرها) أو تسجيلات صوتية (قراءات جهرية لنصوص أدبية أومحادثات بلغة أجنبية أومناظرات) أو تسجيلات فيديو (أداء تمثيلي أو موسيقي أو رياضي). وقد يكون الطالب هو المسؤول عن إغناء البورتفوليو الخاص به بالمواد وتقديمه للمدرس عندما يطلبه هذا الأخير، أو قد يكون المدرس هو المسؤول عن تجميع أعمال الطالب في بورتفوليو واحد جاهز للتقييم في الوقت المناسب. ويتم تقييم البورتفوليو من قبل المدرسين أو الطلاب أنفسهم، وتعطى درجات مناسبة لكل طالب. ويمكن للطلاب وأولياء أمورهم وإداريي المدرسة وغيرهم من المتخصصين المعنيين بتقدم الطلاب الأكاديمي الإطلاع على درجات الطلاب والملفات المختلفة.

تعتبر البورتفوليو نوعاً من الأداء وعليه فإن أي تقييم له هو تقييم للأداء. ومن الضروري جداً توضيح الأغراض التي ترمي إليها البورتفوليو ومعايير تقييمها واستخدامات نتائجها للطلاب وأولياء أمورهم بشكل شفاف وجلي.

تختلف ملفات الإنجاز الشخصي (البورتفوليو) من حيث الوضوح والشمولية والفائدة. ومن بين المعايير التي تم اقتراحها لتقييم البورتفوليو ما اقترحه بولسون وبولسون (١٩٩٢) بتقسيم البورتفوليو بعد تقييمها إلى: متميزة أو ملتزمة بالهدف، أو تحوي بذور تحقيق الهدف، خارجة عن المسار الصحيح. ويضم الشكل ٥-٤ وصفاً دقيقاً للمعايير التي تقسم البورتفوليو على أساسها إلى

الفئات الآنفة الذكر. وكما يبدو من الشكل فإنه يمكن القول بأن البورتفوليو المعدة بشكل جيد تحكي، بشكل أو بآخر، قصة مترابطة عن تطور الأفراد كطلاب علم، إذ أنها توضح درجة الإنجاز الذي حققه الطالب أو التقدم الذي أحرزه، بالإضافة إلى الكشف عن ما يراه الطالب مهماً وجديراً بضمه إلى البورتفوليو وعن طبيعة الطالب كإنسان ذو خصوصية. ولذلك فإن عملية اختيار الطالب للمواد التي يريد أن يضمها البورتفوليو هي دليل واضح على التفكير الذي قام به الطالب لإتخاذ هذه القرارات وعلى مشاركته الفعالة في تحديد ما هو جدير بأن يتم تقييمه.

الشكل ٥-٤

مخطط CMAP للتصحيح الشمولي. (تمت إعادة الطبع بموافقة ف. ل. بولسون و ب. ر. بولسون من كتاب تهيئة البورتفوليو The Making of a Portfolio. حقوق الطبع ١٩٩٢ محفوظة لبولسون وبولسون)

تعريف البورتفوليو

يجيب بورتفوليو الطالب على السؤال الهام "من أنا" ويروي قصةً مترابطة عن الطريقة التي يتعلم بها الطالب، فالبورتفوليوهو مجموعة متكاملة وذات معنى من أعمال الطالب التي تلقي الضوء على جهوده وإنجازاته والتقدم الذي أحرزه في مادة معينة أو عدة مواد. ويتضمن البورتفوليو كذلك دلائل على تقييم الطالب لنفسه ومراجعة حساباته ومساهمته في تسليط الضوء على المهارات التي يراها جديرة بالتقييم من خلال مساهمته الفعالة في اختيار ما يدخل ضمن البورتفوليو. وباختصار شديد، يعد البورتفوليو وسيلة يتواصل بها الطالب مع العالم الخارجي ويعبر من خلالها عن المعلومات والمهارات التي تعلمها وعن أهمية هذه المعلومات والمهارات.

البورتفوليو المتميز

يروي البورتفوليو المتميز قصةً متماسكة وإنسيابية عن ممارسات الطالب كفردٍ مفكر وناقدٍ للذات، إذ يضم أقساماً مختلفة ذات إرتباطات واضحة وتهدف إلى تدعيم هدف رئيسي واحد. كما يمكن للمطّلع على البورتفوليو أن يرى وبشكلٍ واضح الأسباب التي تم على أساسها تقييم المكونات المختلفة له والتي لم تتح مجالاً كبيراً للإختلاف بين كل من ساهم في تقييم عمل هذا الطالب. ويمكّن البورتفوليو المتميز أي مطلّع خارجي من أن يكوّن فكرةً جيدة عن الطالب الذي صنع البورتفوليو وعن الخطوات التي ارتقى على أساسها المستوى التعلّمي لهذا الطالب.

البورتفوليوالذي يسير نحو الهدف

إن البورتفوليوالذي يسير نحو الهدف الذي في طريقه إلى التعبير عن قصة الطالب كطالب علم مستقل. فعلى الرغم من أن نقد وتحليل الذات واضحين في البورتفوليو، إلا أن العلاقات بين العناصر المكونة له ليست واضحة تماماً في بعض الأحيان، كما أن دور المساهمين الآخرين في اختيار عناصره (كالمدرسين مثلاً) ليس جلياً. وعلى الرغم من أن البورتفوليو يضم دلائل على نقد الذات والتفكير التحليلي، إلا أن الأسباب التي جعلت الطالب يختار عناصر البورتفوليو لم يتم توثيقها أو التعبير عنها بالشكل المطلوب. كما يتطلب هذا النوع من الملفات من المراقب الخارجي أن يستنتج بعض صفات الطالب ومدى التطور العلمي الذي أحرزه.

البورتفوليوالذي يضم بذور تحقيق الهدف

يمكن تحديد بعض العلاقات بين العناصر المكونة للبورتفوليو من هذا النوع، إلا أن اختيارات الطالب تبدو وكأنها تمت على أسس حدسية أتبعها الطالب بعد

ذلك بتحليلات جيدة عن العملية التعليمية التي مر بها. وتبقى المشكلة الرئيسة التي يعاني منها هذا النوع من البورتفوليو هي عدم قدرة المراقب الخارجي للبورتفوليوعلى تكوين صورة واضحة عن الطالب أو عن تحديد معالم قصة تطوره العلمي. وبالرغم من وجود أدلة على نقد الذات والتفكير التحليلي يبقى البورتفوليو في مرحلة "البذور" والتي يمكن أن يتجاوزها مع الوقت.

البورتفوليو الذي خرج عن المسار (أو ضل الطريق)

يعتبر هذا النوع من البورتفوليو "حاوية" ليس إلا تضم عينة عشوائية من عمل الطالب لا تنم عن بذل أي جهد تنظيمي من قبل المتعلم للربط بين عناصر هذه العينة. ولا تعطي هذه العينة صورة واضحة عن التحصيل العلمي للطالب كما أنها تعطي إنطباعاً بأن اختيار الطالب لعناصر البورتفوليو كان مبنياً على "ما يريده المدرس" وليس على أي اعتبارات أخرى مما يعني أن القرارات الأساسية في اختيار عناصر البورتفوليو كان مصدرها طرف ثانوي (وليس الطرف الأولي وهو الطالب) أملى على الطالب الاختيارات وحول دور الطالب إلى مجرد منفذ للتعليمات. ويتسم هذا النوع من البورتفوليو بالركود وعدم التعبير عن أي تطور نحو إعادته إلى المسار الصحيح، كما تندر الأدلة على نقد الذات والتفكير التحليلي ولا يحاول الطالب تبرير اختياراته لعناصر البورتفوليو أو تفسير ما قد تعلمه.

وعلى الرغم من توفر معظم تعليمات تجميع البورتفوليو وأمثلة عنها في مقالات المجلات التخصيصة (هربرت، ١٩٩٢؛ وولف، ١٩٨٩) أو في الكتب (كاري، ١٩٩٤؛ ماكدانيال، ١٩٩٤؛ بولسون وبولسون، ١٩٩٢؛ المجلس التعليمي لولاية فيرمونت، ١٩٩١)، تتوفر كذلك عدد من الاختبارات المسوقة تجارياً كبورتفوليو تقييم المهارات اللغوية (خدمة الإرشاد الأمريكية) وبرنامج البورتفوليو الشامل للأدب ومهارات اللغة (شركة ريفرسايد للنشر) وتقييم CTB للبورتفوليو

(ماكغرو – هيل/ CTB). ويتألف بورتفوليو تقييم المهارات اللغوية من ملف جامع لأداء الطلاب وأعمالهم المختلفة للصفوف ١ إلى ٦ في مجالات القراءة والكتابة والإستماع والتحدث والتي قد يقيمها المدرس أوالطالب. كما يتألف برنامج البورتفوليو الشامل للأدب ومهارات اللغة من عدد من وسائل التقييم التي تعد عند الطلب، ويتكون معظمها من أسئلة ذات إجابات قصيرة أو إجابات يمكن اختيارها. وترافق الاختبارات قوائم للرصد لملاحظات المراقبين لأداء الطلاب. وتساعد هذه القوائم على تقييم اهتمامات وتجارب الطلاب (بين الصفوف ٢ و٨) الشخصية في مجالات القراءة وغيرها من المهارات اللغوية. ويتألف تقييم CTB للبورتفوليو من ٨ إلى ٢٠ تجربة مستمدة من الحياة العملية تم تصميمها بهدف تحفيز الطالب على التخطيط والتوصل إلى إجابات إبداعية على مدار السنة الدراسية. ويراقب أولياء الأمور والمدرسون والطلاب أنفسهم تطور المفاهيم والمهارات لديهم.

التقييم الديناميكي

تعد تقييمات الأداء أحد أفضل أنواع التقييم التي تخدم أهدافاً تعليمية وتقييمية (للطلاب) في آنٍ واحد. إذ أن عملية البحث عن إجابات للأسئلة وحلول للمشكلات وإعداد وتنفيذ مشاريع مختلفة وجمع وتحليل ودمج المعلومات أو المواد ومشاركة الآخرين في النقاشات وابتداع الحلول أو المنتجات التي تتوج جهود الطالب تعني أن الطالب يعيش في حالة من التعلم المستمر. ولا يمكن القول أن التعلم هو حكر على تقييمات الأداء إذ يتعلم الأفراد من شتى أنواع التجارب التي يمرون بها بما فيها الاختبارات. ومما لا شك فيه أن الإجابة على أسئلة متعددة الاختيارات والتحضير للاختبارات الشفهية والكتابية يساهمان في إغناء تجارب الطالب. إلا أن اختبارات الأداء تبقى مميزة من حيث كونها اختبارات غير مقيدة، تستند إلى الواقع العملي وتتطلب التزاماً وانغماساً كبيرين من الطالب وتتيح له

فرصة الحصول على تعلّم عملي ومتنوع ويسهل نقله واستعماله خارج حدود المؤسسة التعليمية.

وتعتبر إجراءات التقييم الديناميكية تعبيراً عن قناعة العاملين في مجال التعليم بالإرتباط الوثيق بين التعليم والتعلم والتقييم. ويركز التقييم الديناميكي، شأنه كشأن تقييم الأداء، وبناءً على إجراءات فايغوتسكي (١٩٦٢، ١٩٧٨) لتقييم مستوى التطور العقلي للطفل، على عمليات ونواتج السلوك البشري. وإن الغرض من التقييم الديناميكي لا يتوقف عند حدود تقييم تحصيل الطفل عند نقطة زمنية معينة بل يتعدى ذلك ليقيم الإمكانية الإنجازية عند الطفل وذلك من خلال ملاحظة قدرته على التوصل إلى حلول للمشكلات مع بعض المساعدة. ويمكن استخدام عدد من التقنيات والإستراتيجيات في التقييم الديناميكي كالتدريب والإعداد (coaching) (بدوف، ١٩٨٧)، والوساطة (mediation) (فيورشتاين وراند وجينسين وكانيل وتزورييل، ١٩٨٧)، والتلقين التدريجي qiaduated prompting (كامبيوني وبراون، ١٩٨٧). وتعتمد طريقتا التدريب والوساطة على دورة الاختبار – فالتدريس – فالاختبار. إذ يتم اختبار الطالب ثم إطلاعه على مادة الاختبار وتدريبه عليها ثم اختباره مرة ثانية، ويشير الفرق بين نتيجة الطالب قبل التدريب ونتيجته بعده إلى قدرة هذا الطالب على التعلم (أو ما يسمى بمنطقة التطور الأقرب zone of proximal development). وتختلف طريقة الوساطة عن طريقة التدريب بكون الأولى تعتمد على تدخلات مستمرة من الفرد البالغ لزيادة فرص الطفل في إنجاز المهمة على أتم وجه. أما في طريقة التلقين التدريجي graduated prompting فيقوم القائم على الاختبار بإعطاء الممتحَن تلميحات وإشارات موجودة في نص تم وضعه سابقاً وذلك ليساعد الممتحَن في إنجاز المهمة بنجاح. وتتدرج هذه التلميحات من العام إلى الدقيق والمحدد حسب الحاجة.

الخلاصة

تختلف اختبارات الأداء (موضوع هذا الفصل) عن غيرها من الاختبارات في كونها لا تتوقف عند تقييم معرفة الطالب بل تتطلب من الطالب لا أن يعرف فقط بل أن يصنع وينتج أو يحول معرفته النظرية إلى أشياء عملية. وقد استخدمت اختبارات الأداء التقليدية بشكل أساسي في مجالات التربية البدنية والفنون التطبيقية واليدوية واللغات الأجنبية، بالإضافة إلى القراءة الشفهية والإنشاد والخط (في المرحلة الإبتدائية). وقد تم استخدام الاختبارات الشفهية والمقالية (والتي تعتبر اختبارات أداء للقدرة لتقيّم مدى معرفة الفرد وقدراته الاستنتاجية بشكل صحيح ومقنع إما شفهياً أو كتابياً) في الجامعات والمدارس لمئات السنين. وقد كانت هذه الاختبارات وحتى بدايات القرن العشرين الوسيلة الأساسية لتقييم درجة تحقيق الطالب للأهداف التعليمية المحددة له. إلا أن الإتجاه في القرن العشرين نحو الاختبارات ذات أسئلة الخيارات المتعددة وغيرها من الاختبارات التي يعدها المدرسون أو الاختبارات التحصيلية القياسية قاد إلى إنحسارٍ مستمر في استخدام الاختبارات الشفهية والمقالية.

الاختبار المحكم التصميم، بحسب رأي رواد القياس النفسي والتعليمي أمثال ألفرد بينيه وترومان كيلي و إ. ف. ليندكويست، هو اختبار مؤلف من عينة نموذجية من الفقرات المتعلقة بالمهارات والمعرفة المراد تقييمها. وينطبق هذا على الخيارات المتعددة من متعدد واختبارات الأداء، إلا أن نقاد الاختبارات القياسية يرون أدوات القياس هذه كأدوات غير عملية وغير واقعية تؤكد على التعلم السطحي وتغفل القدرات العميقة والتطبيقية. ولهذا السبب شهدت السنوات الأخيرة إنتعاشاً كبيراً في استخدام الاختبارات الكتابية والمقالية وغيرها من طرق تقييم الأداء. إن معظم هذه الاختبارات أدوات قياس غير منشورة يضعها المدرسون لاختبار طلابهم، كما يتوفرعدد من الاختبارات المنشورة في مجالات

الرياضيات واللغة والعلوم والدراسات الاجتماعية. وتزعم اختبارات الأداء الجديدة هذه بأنها تقيس القدرات التعليمية المعقدة والأكثر شمولاً، كما أنها تركز على القدرات الحياتية العملية. إضافة إلى ذلك، تتسم اختبارات الأداء بأنها تدفع الطلاب إلى العمل (على عكس الاختبارات الموضوعية) وتساهم في خلق فرص للعمل الجماعي بين الطلاب والمدرسين وضمن مجموعات صغيرة من الطلاب. تقيس اختبارات الأداء، ومع الأسف، القدرات التي أغفلتها اختبارات الخيارات المتعددة على حساب تقديم عينات من تجارب تعليمية متعددة وعلى حساب كفاءة الإجراء وموضوعية التصحيح وثبات الاختبار. وهذا لا يعني أن الاختبارات الموضوعية والاختبارات التحصيلية خالية من العيوب. ويعد استخدام نماذج اختبارية تجمع بين هذين النوعين من الاختبارات المعيار الأمثل في السنوات القادمة.

يُستخدم نظام البورتفوليو إلى جانب اختبارات الأداء في تقييم المهارات الأساسية وغيرها من المواد التعليمية. ويقوم نظام البورتفوليو على تجميع عينات من أعمال الطالب (كالمشاريع والإنجازات المدرسية) وتقييمها لاحقاً. ويمكن جمع عدد من الأعمال التي تعبر عن تحصيل الطالب في مادة أو مقرر واحد أو في عدد من المقررات. كما يمكن أن يكون القرار بشأن الأعمال التي يضمها البورتفوليو والمحكات التي سيقيم على أساسها قراراً يتوصل إليه المدرس والطالب سويةً.

إن صعوبة التمييز بين اختبارات الأداء والتدريس الجيد دليل على التأكيد في السنوات الأخيرة على إلتحام عمليتي التقييم والتعليم. ولعل أحد الأمثلة على هذا التحرك في المجال التعليمي عملية التقييم الديناميكي بتأكيدها على الاختبار –فالتعليم– فالاختبار عند تقدير إمكانيات الطلاب.

مصطلحات للمراجعة

راجع معاني المصطلحات المدرجة أدناه والتي تم استخدامها في هذا الفصل. الرجاء مراجعة فهرس التعريفات أو المعجم في حال وجود أي التباس حول معاني المصطلحات.

التقييم البديل	الوساطة
طريقة وضع الدرجات التحليلية	اختبار مهارات الاستماع والحديث
التقييم الواقعي	اختبار التحصيل الشفهي
التدريب والإعداد	اختبار الأداء
معامل المطابقة	بورتفوليو (ملف الإنجاز الشخصي)
التقييم الديناميكي	قواعد التقييم rubrics
استراتيجية التلقين التدريجي	تصنيفات السلوك النفسي – الحركي
طريقة وضع الدرجات الشمولية	قائمة رصد لتقييم أداء المجموعات
ثبات الاختبار مع اختلاف المصحيحن	منطقة التطور الأقرب
معامل الارتباط الصفي	

المراجع المقترحة

Baker, E. L. (1994). Making performance assessment work: The road ahead. *Educational Leadership, 51*(6), 58–62.

Bateson, D. (1994). Psychometric and philosophic problems in "authentic" assessment: Performance tasks and portfolios. *Alberta Journal of Educational Research, 40*(2), 233–245.

Coleman, L. J. (1994). Portfolio assessment: A key to identifying hidden talent and empowering teachers of young children: Javits grant projects [Special issue]. *Gifted Child Quarterly, 38*(2), 65–69.

Guskey, T. R. (1994). What you assess may not be what you get. *Educational Leadership, 51*(6), 51–54.

Hart, D. (1994). *Authentic assessment: A handbook for educators.* Menlo Park, CA: Addison-Wesley.

Herbert, E. A. (1992). Portfolios invite reflection—from both students and staff. *Educational Leadership, 49*(8), 58–61.

LeMahieu, P. G., Gitomer, D. H., & Eresh, J. (1995). Portfolios in large-scale assessment: Difficult but not impossible. *Educational Measurement: Issues and Practice, 14*(3), 11–16.

Linn, R., Baker, E., & Dunbar, S. (1991). Complex performance based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15–21.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher, 23*(2), 13–23.

Newmann, F. M. (1997). Authentic assessment in social studies: Standards and examples. In G. D. Phye (Ed.), *Handbook of classroom assessment: Learning, achievement, and adjustment* (pp. 359–380). San Diego, CA: Academic Press.

Smith, J. K. (1994). Standardized testing versus authentic assessment: Godzilla meets Winnie-the-Pooh. In L. M. Morrow, J. K. Smith, & L. C. Wilkinson (Eds.), *Integrated language arts: Controversy to consensus* (pp. 215–229). Boston: Allyn & Bacon.

Stecher, B. M., & Herman, J. L. (1997). Using portfolios for large-scale assessment. In G. D. Phye (Ed.), *Handbook of classroom assessment: Learning, achievement and adjustment* (pp. 491–516). San Diego, CA: Academic Press.

٦

اختبارات الاستعداد الواحد والاستعدادات المتعددة

تشير كلمة تحصيل Achievement إلى ما قد تم التوصل إليه أو إنجازه كقراءة كتاب ما أو القدرة على التحدث بلغة أجنبية معينة أو فهم كيفية حل معادلة جبرية معقدة. أما كلمة الاستعداد Aptitude فتشير إلى ما يمكن لفردٍ ما إنجازه في حال توفر الظروف التي تساعده على ذلك، وبعد تلقيه للتدريب أو الخبرات المناسبة. ويشمل مصطلح القدرة Ability مصطلحي التحصيل والاستعداد إذ أن ما يقدر عليه المرء هو مجموع ما قد حققه أو توصل إليه وما هو قادرٌ على تعلمه.

وتتشابه مقاييس التحصيل ومقاييس الاستعداد من حيث التركيب فيصعب على المطلع عليها تحديد الاختلافات بينها من النظرة الأولى فكلا النوعان من الاختبارات يقيسان مهارات الفرد ومعرفته في مرحلة معينة من تطوره. إلا أنه وعلى حين تركز الاختبارات التحصيلية على معلومات محددة اكتسبها الطالب من مقرر دراسي معين، تقيس اختبارات الاستعداد المهارات والكفاءات التي اكتسبها الفرد من المصادر التعليمية العامة وغير الرسمية التي توفرها الحياة العملية خارج حدود المدرسة أو المؤسسة التعليمية.

ومن المريح لنا هنا أن ننظر إلى اختبارات التحصيل والاستعداد على أنها أجزاء من حقلٍ تعليمي يضم خبرات شديدة الخصوصية وأخرى عامة، وتتوضع الاختبارات التي تقيس الخبرات الإنسانية العامة كالاختبارات غير اللغوية واختبارات الأداء والاختبارات العادلة حضارياً التي تقيس القدرات الإدراكية للإستجابات الفطرية (كاختبار أوقات ردود الأفعال) في موقع أقرب إلى الخبرات العامة منها إلى الخبرات الخاصة ضمن الحقل التعليمي المذكور أعلاه. بينما تقع الاختبارات التي تقيس التحصيل في مقررات دراسية معينة أو مهارة عملية ما في موقع أقرب إلى الخبرات الشديد الخصوصية من الحقل التعليمي. أما اختبارات التحصيل التي تقيس المهارات العامة واختبارات الذكاء والاستعداد اللغوي فتحتل موقعاً متوسطاً من هذا الحقل (أنستازي وأوربينا، ١٩٩٧).

لقد كان الإعتقاد السائد بين العاملين في مجال الاختبارات العقلية هو أن الوراثة هي المسؤول الوحيد عن تحديد قابليات الفرد الخاصة وقدراته الذهنية العامة أو درجة ذكائه، إلا أنهم وعلى الرغم من هذا الاعتقاد كانوا على قناعة أنه لا يمكن قياس قابليات الفرد بشكلٍ مباشر، مما يحتم اللجوء إلى طرق قياس غير مباشرة تقيس ما قد تعلمه الفرد بالصدفة أو بشكلٍ غير رسمي وغير موجه خارج حدود المؤسسات التعليمية. وعليه فإن إدراج مصمم الاختبار لأسئلة عامة ضمن اختباره تهدف إلى تقييم التجارب التي يتعرض إليها جميع البشر دون إستثناء، سوف يتيح الفرصة للأفراد الذين تعلموا أكثر من هذه التجارب بأن يجيبوا بشكل أفضل على أسئلة الاختبار.

تُستخدم اختبارات الاستعداد بشكلٍ أساسي للتنبؤ بأداء فردٍ ما إذا ما أتيحت له الفرصة بتعلم مهارة ما إما ضمن إطار مقرر أكاديمي أو برنامج تدريبي أو أثناء ممارسة مهنة معينة. إلا أنه لا يمكن القول بأن الاختبارات التحصيلية لا تؤدي المهمة ذاتها وبنفس الدرجة من الكفاءة مما يعني أن الفرق

بين هذين من النوعين من الاختبارات مبرر إلا أنه غير واضح على الإطلاق. ويمكن العودة إلى الدراسة التي قام بها كارول (١٩٧٣) للإطلاع على تبريرات لضرورة التمييز بين اختبارات التحصيل واختبارات الاستعداد. فعلى سبيل المثال، توافقت درجات تحصيل عدد من الطلاب في مقرر لغة أجنبية معينة لم يكن أي من الطلاب قد تعلمها قبل إلتحاقهم بهذا المقرر مع الدرجات التي تنبأ بها اختبار الاستعداد الذي أجري لهؤلاء الطلاب. وكما كان متوقعاً فإن درجات الطلاب في اختبار التحصيل ارتفعت مع تقدم الطلاب في دراسة المقرر على حين لم يطرأ أي تغيير على درجات اختبار الاستعداد عند إعادة الاختبار.

وتعزى بعض الإختلافات بين اختبارات التحصيل والاستعداد إلى تاريخ نشأة هذين النوعين من الاختبارات. ففي حين تعود جذور الاختبارات التحصيلية إلى المؤسسات التعليمية التي استخدمت هذه الاختبارات كأدوات قياس تكوينية وتراكمية لتحديد تقدم الطالب في المقررات المختلفة، وُلدت اختبارات الاستعداد في الشركات والمؤسسات المهنية حيث استخدمت هذه الاختبارات في اختيار الموظفين وتوزيعهم على الأقسام والمديريات المختلفة وفي التوجيه والإرشاد المهني. وبالرغم من الارتباط التاريخي بين الاختبارات التحصيلية والمدارس، تم تصميم عدد كبير من هذه الاختبارات للاستخدام في مجال الأعمال والتوظيف. وينطبق هذا الأمر كذلك على اختبارات الاستعداد التي تم تصميم عدد كبير منها للاستخدام في المدارس والجامعات. ومن الجدير بالذكر أن عدداً كبيراً من الاختبارات التحصيلية واختبارات الاستعداد يُستخدم في مجالات التعليم والحرف والجيش والدولة وغيرها من المجالات.

وكما هو الحال في أنواع أخرى كثيرة من الاختبارات النفسية فإن حقبة العشرينيات من القرن المنصرم تعد الحقبة التي شهدت التطور الأكبر في اختبارات الاستعداد. وقد كان هذا التطور نتيجةً حتمية للرغبة في زيادة الإنتاج

ورفع الكفاءة التي كانت الشغل الشاغل للكثيرين في ذلك الوقت. كما شهدت هذه الحقبة ثلاث محاولات رائدة لمعالجة هذه المشكلة كان أولها تلك التي قام بها فريدريك تايلور لإعادة تصميم أدوات ومعدات العمل وتخفيف أعباء العمال بغرض رفع الكفاءة. أما المحاولة الثانية فيعود الفضل فيها إلى فرانك وليليان غيلبرث وأبحاثهما المعروفة بأبحاث "الوقت والحركة" والتي استهدفت إعادة تصميم متطلبات الوظائف المختلفة. أما المحاولة الثالثة فقد هدفت إلى تعزيز الكفاءة الإنتاجية عن طريق اختيار العمال القادرين على أداء المهمات الموكلة إليهم بأعلى درجات الكفاءة وأقل درجات هدر الوقت والجهد.

وقد دفع النجاح الكبير الذي حققه علماء النفس في قياس قابليات المجندين في الحرب العالمية الأولى إلى عرض خدماتهم على المؤسسات والشركات لمساعدتهم على اختيار موظفيهم بعد إخضاعهم لاختبارات استعداد تحدد مدى مناسبتهم للوظائف التي تقدموا لها. وقد أدت الحاجة الكبيرة آنذاك إلى عمال الآلات وتقنيي الصيانة وموظفي المكاتب إلى قفزات هائلة في مجال اختبارات القابليات النفسية - الحركية والمكتبية والميكانيكية. ولعل أهم الأبحاث العلمية في تلك الفترة هو البحث الذي قام به د. ج. باترسون ومساعدوه في جامعة مينيسوتا (باترسون وأليوت وأندرسون وتوكس وهايدبردر، ١٩٣٠) والذي قاد إلى وضع عدد من الاختبارات الميكانيكية والمكانية والمكتبية. كما تم وضع عدد من الاختبارات النفسية - الحركية كاختبار جونسون أوكونور للمهارات الحركية البسيطة. وتم إعداد عدد من اختبارات القابليات لأمور أقل مهنية كاختبار نورمان ماير للفنون واختبار كارل سيشور للاستعداد الموسيقي. كان استخدام هذين الاختبارين لأغراض البحث وليس التوظيف.

وشهدت بدايات الثلاثينيات ظهور نوع بدائي من بطاريات اختبارات الاستعداد العامة وهي مجموعة من اختبارات القدرة لا تزال مستخدمة حتى اليوم في مجال الإرشاد الوظيفي وبرامج التنسيب المهني بواسطة خدمات التوظيف

الأمريكية. كما شهد عام ١٩٤١ حدثاً يعد من معالم تاريخ تطور الاختبارات وهو تطوير بطارية اختبارات تصنيف المتطوعين في سلاح الجو والذي يقيس المهارات المطلوبة للعمل في مجال الطيران. وقد ساهم هذا الاختبار في تفعيل عملية اختيار الطيارين والملاحين وقاذفي القنابل مما أدى (بعد الحرب العالمية الثانية) إلى العمل على تطوير بطاريات أخرى لاختبارات الاستعداد.

يهدف هذا الفصل من الكتاب إلى إلقاء الضوء على اختبارات الاستعداد واستخداماتها في إتخاذ القرارات المهنية في مؤسسات الأعمال والصناعة والجيش والدولة والقرارات التعليمية في المدارس والجامعات. وعلى حين صممت بعض هذه الاختبارات لقياس استعداد واحد كالاستعداد للأعمال المكتبية أو الميكانيكية أو برمجة الحاسب أو تشغيل الآلات أو غيرها، تم تصميم بعض اختبارات الاستعداد لقياس قابليات متعددة كبطارية اختبارات الاستعداد العامة واختبارات القابليات التباينية وبطارية القابليات المهنية لعناصر القوات المسلحة. وتتمتع الاختبارات التي تقيس استعداد واحد بنطاق ضيق وبدقة وأمانة (fidelity) أعلى في القياس من الاختبارات التي تقيس عدة قابليات أما الاختبارات واسعة النطاق فتشير إلى مدى المحكات المتوقعة من الاختبار والقرب من الصحة أو مدى الصدقية التي بواسطتها يتم التنبؤ بالمحكات.

تعد مؤسسات الجيش والقوات المسلحة الجهة الأكثر استخداماً لاختبارات الاستعداد (التي تقيس كل استعداد على حدة أو عدداً من القابليات في آنٍ واحد)، إذ يتم إخضاع أكثر من ١,٥ مليون شخص كل عام لبطارية اختبار القابليات المهنية لعناصر القوات المسلحة الأمريكية ويستخدم إلى جانبه أكثر من ٩٠٠ اختباراً للمهارات المختلفة. وتستخدم قطاعات الأعمال الخاصة والصناعة والمؤسسات الحكومية والفدرالية والمحلية عدداً من اختبارات الاستعداد التي تقيس المهارات المكتبية إلا أن عدد الاختبارات التي تستخدمها جميع هذه

المؤسسات مجتمعة لا يقترب من العدد الإجمالي للاختبارات المستخدمة من قبل الجيش في الولايات المتحدة الأمريكية. وبالرغم من أن كثير من هذه الاختبارات هي اختبارات تحصيلية أو كفاءة تقيس القدرة على إنجاز مهمات عسكرية معينة، إلا أن أغلب الاختبارات المستخدمة في الجيش هي اختبارات تقيس استعداد الفرد لتعلم أعمال وعمليات محددة.

ويغض النظر عن المجال الذي يراد استخدام اختبارات الاستعداد للتنبؤ بالأداء والسلوك فيه، يجب أن يكون مستخدم الاختبارات على دراية كاملة بالإحصائيات عن دقة التنبؤ وبالتداعيات القانونية التي يمكن أن تنجم عن استخدام نتائج اختبارات الاستعداد في عمليتي الاختيار والقبول للموظفين والطلاب ضمن مجالات التعليم والتوظيف. وقد تم التطرق إلى هذه الأمور في الفصلين الأول والثاني من الكتاب وسوف نتحدث عنها في هذا الفصل أيضاً بعد إعطاء فكرة عامة عن أهم اختبارات وبطاريات الاستعداد المستخدمة.

اختبارات الاستعداد الواحد

لا يعتمد أداء الفرد في الاختبارات النفسية والتعليمية على خصائص إدراكية وعاطفية وحسية -حركية معقدة فقط، بل يعتمد كذلك على قدرات البصر والسمع والحركة والسرعة والدقة وغيرها من القدرات الحسية - الحركية. ولذلك، وكما يملي قانون الدولة، من المهم جداً أن يتم فحص القدرات الحسية - الحركية للأطفال والبالغين بشكل دوري إذ أن وجود خلل بسيط في هذه القدرات قد لا يكون من السهل التعرف عليه بمجرد النظر إلى الشخص أو التحدث معه. كما أنه من الممكن أن يكون الشخص نفسه لا يعلم بوجود هذا الخلل نظراً لضآلته إلا أنه يمكن أن يؤثر على أدائه في اختبار القدرة مما قد يعني عدم حصوله على الوظيفة التي تقدم لها أو القبول في المدرسة أو الجامعة التي أراد الانتساب إليها. والجدير بالذكر أن التعليمات اليوم تؤكد على حتمية

قبول ذوي الإعاقات في الوظائف والمؤسسات التعليمية المختلفة وتأمين الظروف المواتية للتغلب على إعاقاتهم.

يمكن التأكد من الحدة السمعية والبصرية لفردٍ ما عن طريق استخدام إجراءات بسيطة (مثل لوحة سنيلين واختبارات الساعة[١] وغيرها) يمكن للمدرس أو المساعد أن يقوم بها أو إجراءات معقدة (استخدام منظار العين أو مقياس السمع) تتطلب الإستعانة بطبيب عيون أو أخصائي فحص العيون أو اخصائي فحص السمع. ويتضمن الاختبار الشامل لحدة البصر اختبارات للقدرة البصرية في كل عين وللعينين معاً ولرؤية القريب والبعيد وللتوازن العضلي لكل عين أثناء النظر إلى الأشياء القريبة والبعيدة ولإدراك العمق ولدقة رؤية الألوان. وقد تم استخدام عدد من أدوات القياس لتقييم القدرة البصرية للأفراد كاختبار B&L للبصر والذي يُستخدم غالباً في مجالات الصناعة لتصنيف الأفراد ويتم تقييم نتائج الاختبار بالأداة بناء على عدد مجموعات العمل البصرية وذلك بالاعتماد على القدرات البصرية الأساسية لمهنة معينة.

يتم اختبار قدرة الأفراد على تمييز الألوان عن طريق سلسلة من بطاقات تتألف من ألوان متشابهة تحمل أرقاماً أو أشكالاً مؤلفة من نقاط ملونة وراءها خلفية من نقاط ذات ألوان متعاكسة. ومن الأمثلة على هذه النوع من الاختبارات اختبار دفوارين لحدة البصرفي تمييز اللون. ويضم برنامجا G-1 وG-2 على القرص المرفق بالكتاب اختبارات تقريبية لحدة البصر وتمييز الألوان. بينما يضم برنامج G-5 اختباراً معقداً لسرعة وصحة الإدراك الحسي كالقدرة على تحديد سلسلة من الأرقام في مصفوفة عددية. ويلعب زمن الاستجابة دوراً كبيراً في هذا الاختبار.

---

(١) وهو اختبار لحدة السمع يعتمد على قدرة الأفراد في الاستماع إلى حركة عقارب الساعة والإلتفات إليها على الرغم من محاولات منفذ الاختبار جذب نظرهم بالاتجاه المعاكس. (المترجم).

تحتاج أعداد كبيرة من الحرف كمراقبة وتشغيل جهاز رصد الأمواج الصوتية إلى درجة عالية من الدقة السمعية. ويمكن تقدير الدقة السمعية لفردٍ ما عن طريق إجراء اختبار الساعة إلا أن تحديد هذه الدقة بشكل علمي يتطلب استخدام جهاز مقياس السمع والذي يتم تحويل النتائج التي يولدها إلى مخطط بياني يوضح حساسية كل أذن للنغمات الصافية التي تغطي المجال الكامل للترددات السمعية عند البشر. كما يمكن قياس قدرة الأفراد على تحديد جهة الصوت وعلى التمييز بين محفزات مختلفة النبرة والعلو أو الخفوت (راجع البرنامجين G-3 وG-4).

الاختبارات النفسية -الحركية

يضم برنامج G-6 واحد من أكثر الاختبارات إستخداماً في قياس زمن رد الفعل الحركي ويقيس هذا الاختبار ثلاثة أنواع من أوقات ردود الفعل الحركية للمحفزات البصرية والتي تعرف بأنواع دوندرز (أ) و(ب) و (ج) لزمن ردود الأفعال. وأما النوع (أ) فهو زمن رد الفعل البسيط الذي يولد إستجابة واحدة لمحفّز واحد، والنوع (ب) هو رد الفعل الإختياري الذي يستجيب على أساسه الفرد لمحفز واحد ويهمل الآخر، وأما النوع (ج) فهو رد فعل اختياري تقسيمي يعني إستجابة الفرد برد فعل واحد لمحفز معين ثم الإستجابة لمحفز آخر برد فعل آخر. وبالإضافة إلى تحديد المعدل لزمن ردود الأفعال- للأنواع (أ) و(ب) و(ج)- بالأجزاء الألفية للثانية، يحدد البرنامج الخط القاعدي لزمن رد الفعل وزمن التعرف (على المحفز) وزمن الاختيار (لرد الفعل).

ولا يعتبر زمن رد الفعل أهم عناصر القدرة النفسية – الحركية، إذ تساهم عوامل السرعة والقوة والنشاط في رفع كفاءة الأداء الحركي. وتُستخدم المقاييس التي تقيِّم هذه الخصائص في اختيار العاملين في عدد كبير من المجالات وتعتبر هذه المقاييس أدوات تنبؤ صادقة للأداء في الأعمال المرهقة جسدياً (راجع

بلاكلي، كوينونز، كروفورد وياغو، ١٩٩٤؛ هوغان وكويغلي، ١٩٩٤). وتتوفر اختبارات تتطلب حركات عضلية بسيطة وأخرى تتطلب حركات معقدةتقيس الدقة والثبات الحركي في استخدام الأصابع والأيدي والسواعد والأرجل بالإضافة إلى الاختبارات القوة الإيسومترية[٢].

يقيس اختبار سترومبيرغ للبراعة اليدوية سرعة وصحة حركة الأصابع واليدين والسواعد وقد تم استخدامه لتقييم كفاءة عمال المصابغ وعمال الصب وعمال المكابس وعمال اللحام وعمال التجميع عن طريق اختبارهم في سرعة تنسيق ٥٤ قرص تشبه قطع البسكويت وتأتي في ثلاثة ألوان (أصفر وأزرق وأحمر) ثم إدخالها وفق ترتيب معين داخل لوح خاص وبأقصى سرعة ممكنة (انظر الشكل ٦-١).

ومن اختبارات تقييم مهارات البراعة اليدوية واختبار كروفورد للبراعة في التعامل مع الأجزاء الصغيرة واختبار لوح المسامير ذو الثقوب واختبار لوح بوردو للمسامير[٣]. ويتطلب هذا الأخيران إدخال مسامير في فتحات صغيرة باستخدام كلتا اليدين معاً وكل يد على حدة ثم وضع حلقات فوق المسامير. وتتوفر كذلك اختبارات أخرى مشابهة كاختبار أوكونور للبراعة في استخدام الأصابع واختبار أوكونور للبراعة في استخدام الملقط الصغير.

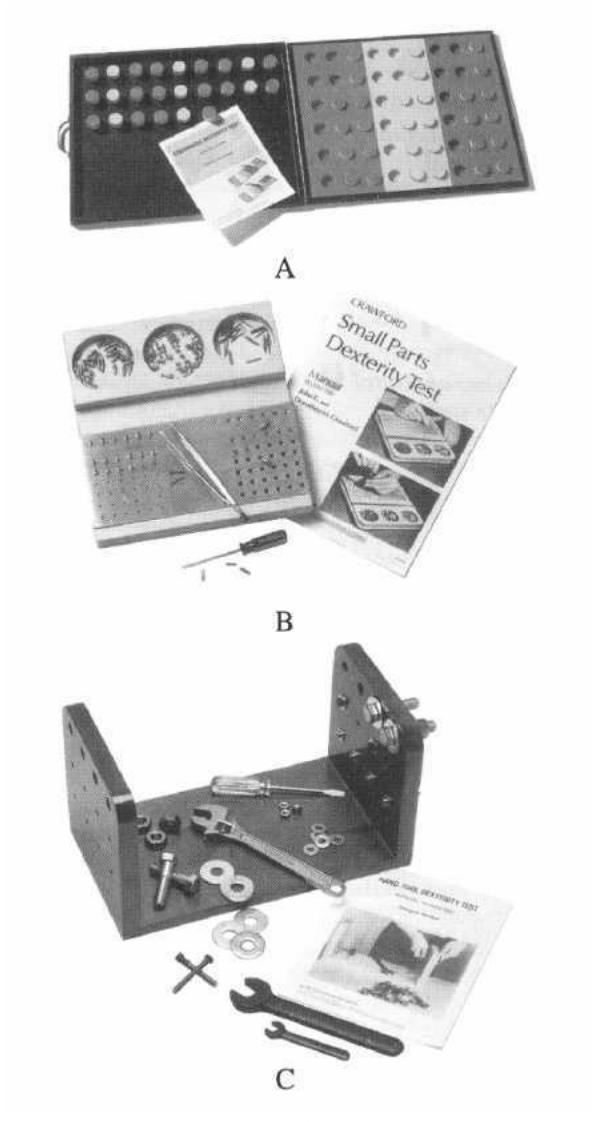
ويقيس اختبارا البراعة في استخدام المعدات اليدوية واختبار مينيسوتا للبراعة اليدوية القدرات على القيام بأعمال يدوية دقيقة اجمالية. يطلب من الممتحن في اختبار البراعة في استخدام المعدات اليدوية يطلب من الممتحَن حل الصمولات حول ١٢ برغي من ثلاثة أحجام مختلفة مثبتة على الجانب الأيسر من لوح خشبي قائم ثم إعادة تركيب البراغي والصمولات على الجانب الأيمن من

(٢) قياس التقلصات العضلية بعد إخضاعها لضغط مقاومة يبقي طول العضلة كما هو. (المترجم).

(٣) يضم كلا اللوحان ثقوب صغيرة يتوجب على الممتحن إدخال المسامير فيها. (المترجم).

اللوح (انظر الشكل ٦-١). ويتألف اختبار مينيسوتا للبراعة اليدوية والذي كان يعرف سابقاً باسم اختبار مينيسوتا لمعدل سرعة المعالجة من جزئين يقيس الأول منهما سرعة الحركة اليدوية للفرد عن طريق اختبار ترتيب يقيس سرعة حركات اليد غير الدقيقة يطلب من الممتحن فيه وضع ستين مكعباً في حفر على لوح كبير، وأما اختبار تدوير فإنه يقيس سرعة حركات الأصابع الدقيقة ويطلب من الممتحَن فيه قلب المكعبات ثم إعادة وضعها على اللوح.

يعد ثبات الاختبارات النفسية - الحركية أقل من غيرها من اختبارات المهارات الخاصة، أما من حيث صدق هذه الاختبارات فتتمتع الاختبارات النفسية - الحركية بارتباط أعلى مع محكات أداء الممتحنين في البرامج التدريبية من اختبارات الكفاءة المهنية. كما تتمتع هذه الاختبارات بارتباط أعلى مع محكات الأداء في أعمال التجميع الرتيب وتشغيل الآلات وغيرها من الأعمال التكرارية من ارتباطها مع محكات الأعمال المعقدة التي تتطلب مستوى أعلى من القدرات الإدراكية والحسية - الإدراكية.



الشكل ٦-١

الأدوات المستخدمة في اختبار سترومبيرغ للبراعة (A) واختبار كروفرد للبراعة في التعامل مع الأجزاء الصغيرة (B) واختبار الدقة في استخدام الأدوات اليدوية (C) (حقوق الطبع محفوظة لمؤسسة علم النفس المتحدة. ١٩٤٥، ١٩٤٦، ١٩٥١، ١٩٥٦، ١٩٦٥، ١٩٨١، تم إعادة الطبع بناءً على موافقة المؤسسة)

القدرات الميكانيكية والمكانية

تُعزى القدرات والمعرفة الميكانيكية التي يملكها العاملون في مجالات الصيانة والآلات وغيرهم من البارعين في فهم وإصلاح وابتكار الأدوات الميكانيكية إلى الخبرة والاهتمام بالإضافة إلى الميول الفطرية. وعلى الرغم من وجود دلائل على ضعف ميكانيكي لدى بعض الأفراد، لا يمكن القول بأن القدرة الميكانيكية هي إمكانية أحادية العنصر، فالمهارة الميكانيكية هي مجموعة من القدرات المعقدة كالسرعة والتنسيق العضلي والإدراك الحسي المكاني والإستيعاب الميكانيكي وغيرها من العوامل الإدراكية. وعلى حين يعد الارتباط بين هذه العناصر للقدرة الميكانيكية ارتباطاً ضعيفاً، يعتبر الإرتباط بين اختبارات القدرة الميكانيكية المختلفة ارتباطاً واضحاً. ومن الجدير بالذكر أن هنالك فرق واضح بين أداء الذكور والإناث في الاختبارات الميكانيكية، إذ غالباً ما يتفوق الذكور في نواحي الإدراك الحسي المكاني والاستيعاب الميكانيكي، وتتفوق الإناث في مهام البراعة اليدوية وبعض عناصر التمييز الإدراكي الحسي. وتصبح هذه الفروقات أكثر وضوحاً مع انتقال الذكور والإناث من مرحلة الطفولة إلى مرحلة البلوغ حيث تساهم العوامل الثقافية والأدوار المختلفة التي يُوجّه الذكور والإناث إليها في إتساع الفجوة في أداء الجنسين في الاختبارات الميكانيكية.

يعد اختبار ستنكويست الذي نُشر عام ١٩١٥ أول اختبار قياس للقدرة الميكانيكية، كما تعتبر الاختبارات التي تمخضت عنها سلسلة الأبحاث في مجال اختبار القدرة الميكانيكية التي أجريت في جامعة مينيسوتا خلال الثلاثينيات من القرن العشرين حدثاً بارزاً في تاريخ تطور أدوات قياس المهارات الميكانيكية. وتقيس هذه الأدوات، وهي: اختبار مينيسوتا للتجميع الميكانيكي واختبار مينيسوتا للعلاقات المكانية واختبار مينيسوتا التحريري للوح الأشكال، البراعة

اليدوية (الاختبار الأول فقط) والإدراك الحسي المكاني (الاختبارات الثلاثة) (باترسون وغيره، ١٩٣٠). وقد خلُصت هذه الاختبارات إلى أن الإدراك الحسي المكاني (وهو القدرة على تصور مواقع الأشياء ضمن فضاء ثلاثي الأبعاد ثم التعامل مع هذه الأشياء للتوصل إلى ترتيب بنائي معين) هو أحد أهم العوامل المؤثرة على القدرة الميكانيكية. تم إعداد اختبار مينيسوتا التحريري (الورق وقلم الرصاص) للوح الأشكال(٤) Minnesota Paper Form Board أول الأمر على أنه النسخة الكتابية المحورة لاختبار مينيسوتا للعلاقات المكانية Minnesota spatial Rolations Test، إلا أن الإرتباط بين درجات الاختبارين كان أدنى من معاملات الثبات للصور المتكافئة.

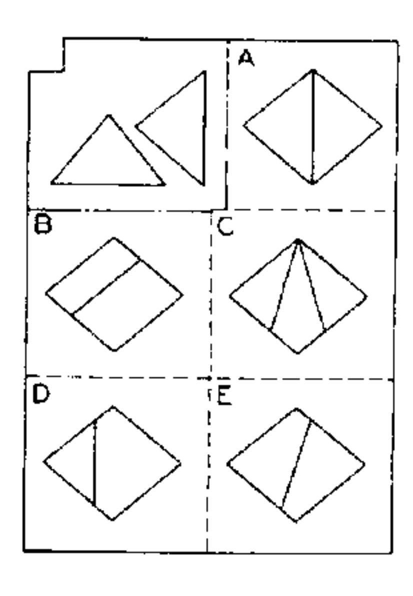
وكما لاحظنا من تجربة برنامج مينيسوتا في مجال اختبارات القدرة الميكانيكية، يمكن استخدام الاختبارات التحريرية واختبارات الأداء لتقييم المهارة الميكانيكية. وتتوفر اليوم نسخ تجارية من اختبارات معدّلة انحدرت من الاختبارات الأصلية التي أعدها برنامج جامعة مينيسوتا للبحث في مجال اختبارات القدرة الميكانيكية، وهذه الاختبارات هي: اختبار مينيسوتا للعلاقات المكانية – النسخة المعدلة (وهو اختبار أداء) واختبار مينيسوتا التحريري المعدل للوح الأشكال (وهو اختبار تحريري). ويتطلب الاختبار الأول من الممتحنين وضع مجموعتين من الأشكال الهندسية الثلاثية الأبعاد في أماكنها المناسبة على أربعة ألواح وفي أسرع وقت ممكن. ويتعين على الممتحن وضع إحدى المجموعتين في اللوحين (أ) و (ب) والمجموعة الثانية في اللوحين (ج) و (د). أما في اختبار مينيسوتا التحريري المعدل للوح الأشكال فيطلب من الممتحَن اختيار الشكل المناسب الذي سوف ينتج عن تجميع عدد من الأجزاء ضمن شكلٍ واحد (راجع الشكل ٦–٢ لمثال على أسئلة هذا الاختبار).

(٤) وهو لوح مشابه للألواح المستخدمة مع الأطفال التي تتطلب من الطفل وضع الأشكال الهندسية في الفتحات المناسبة المحفورة في اللوح. (المترجم).

تتميز درجات المتحنين في اختبار مينيسوتا التحريري المعدل للوح الأشكال بدرجة صدق تنبؤي متوسطة بأداء الأفراد في المقررات والمهن التي تتطلب مهارات مكانية وميكانيكية. فعلى سبيل المثال، تتمتع هذه الدرجات بارتباط ضعيف إلا أنه دال إحصائياً مع درجات الطلاب في مقرارات الهندسة ومع تقييم المشرفين على العمال وسجلات مراقبة الإنتاج ومع أداء العاملين في مجالات التعليب وتشغيل الآلات وغيرها من المهن الصناعية بالإضافة إلى طب الأسنان والفنون.

ويلعب الإستيعاب الميكانيكي إلى جانب المهارات المكانية والنفسية - الحركية دوراً هاماً في تحديد الكفاءة الميكانيكية للأفراد إذ يُفترض في عمال الصيانة والآلات والبناء فهم مبادئ علم الميكانيك. وبالرغم من ارتباط الإستيعاب الميكانيكي بالذكاء العام إلا أن وضع مقاييس خاصة بقياس الفهم الميكانيكي قد أثبتت جدواها. ومن الأمثلة على هذه الاختبارات، اختبار بينيه للإستيعاب الميكانيكي واختبار SRA للاستعداد الميكانيكي. ويتألف اختبار بينيه من ٦٨ فقرة تضم رسومات وأسئلة متعلقة بتطبيق القوانين الفيزيائية والإجراءات الميكانيكية في مواقف عملية، وكما توضح عينة الأسئلة من هذا الاختبار (انظر الشكل ٦-٣) فإن اختبار بينيه يركز على الإدراك الحسي المكاني ومعرفة الأدوات وليس على البراعة اليدوية.

هناك جزءان في الزاوية العليا على اليسار. انظر إلى الأشكال الخمسة المميزة بالأحرف A - B - C - D - E عليك أن تقرر أي من هذه الأشكال يظهر أن هذين الجزئين يمكنهما الانطباق معاً.



الشكل ٦-٢

عينة لسؤال من اختبار مينيسوتا التحريري المعدل للوح الأشكال. (حقوق الطبع محفوظة لمؤسسة علم النفس المتحدة عام ١٩٤١ وجددت عام ١٩٦٩، تمت إعادة الطبع بناءً على موافقة المؤسسة "RMPFBT" وهي علامة مسجلة لمؤسسة علم النفس المتحدة)

يعد اختبار SRA للاستعداد الميكانيكي اختباراً تحصيلياً أكثر منه اختباراً للاستعداد، ويتألف هذا الاختبار من عدد من الاختبارات الفرعية التي تقيِّم المعرفة الميكانيكية والعلاقات المكانية والإستدلال الميكانيكي. وتستخدم نتائج هذا الاختبار (كما كان الحال في اختبار بينيه) في التنبؤ بأداء الملتحقين ببرامج

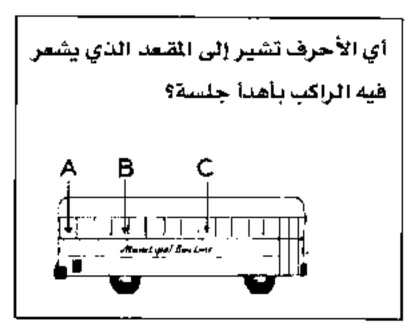
التدريب والوظائف التي تحتل موقع البداية في السلم الوظيفي بالإضافة إلى مجال الإرشاد المهني. وتحدد نسبة الارتباط بين نتائج اختبارات الاستعداد الميكانيكي (كاختبار SRA) ومقاييس أداء عمال الصيانة والآلات في الدورات التدريبية والوظائف بحوالي ٠،٣٠ (غيسيللي، ١٩٧٣). وبشكل عام فإن نسبة الارتباط بين درجات اختبارات القدرة الميكانيكية ومقاييس أداء المتدربين أعلى من نسبة الارتباط بين درجات هذه الاختبارات ومقاييس الكفاءة الوظيفية.

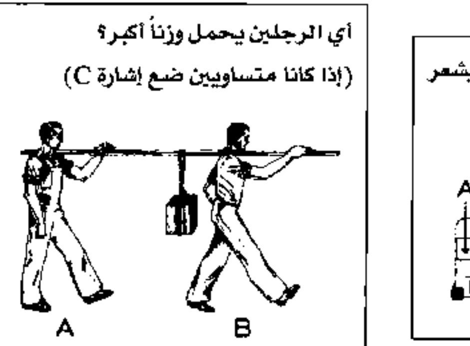
قدرات الوظائف الكتابية

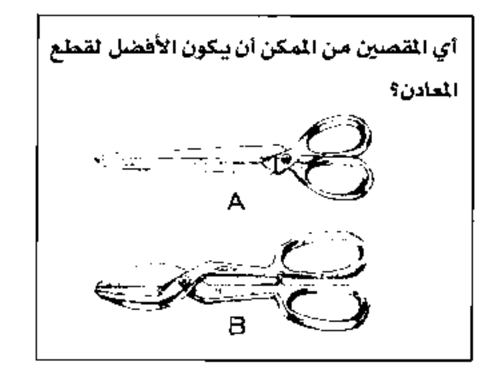
يحتل قطاع الوظائف الكتابية المدنية الدرجة الثانية من حيث عدد الاختبارات التي تُجرى فيه سنوياً، ويأتي بعد القطاع العسكري. ترتبط القدرات على تأدية الوظائف الكتابية على أكمل وجه بعدد كبير من العوامل المعقدة والمتداخلة بالإضافة إلى ارتباطها بالقدرات العقلية العامة. ويحتاج الأداء الجيد في المهام الكتابية للبراعة اليدوية والقدرة على التعرف على أوجه الاختلاف والتشابه بين الأشياء بسرعة، إلا أن المهارات الحسابية واللغوية والتي تدخل في إطار القدرات الدراسية العامة لها التأثير الأكبر على أداء الأفراد في الوظائف الكتابية. ومن أبسط الاختبارات المسوقة تجارياً والتي تقيس القدرات الكتابية اختبار مينيسوتا للمهارات الكتابية (MCT). وتعزى بساطة هذا الاختبار إلى أنه يطلب من الممتحنين تدقيق قوائم تضم أسماء وأرقام فقط. ويُستخدم اختبار MCT في اختيار المتقدمين بطلبات عمل للوظائف الكتابية وفي تقييم أداء الكتّاب وغيرهم من العاملين في المجالات التي تتطلب السرعة الإدراكية الحسية والصحة والقدرة على التعامل مع الرموز (اللغوية والعددية). يحتوي الشكل ٦-٤ على عينة من أسئلة هذا الاختبار، وينقسم الاختبار إلى قسمين: قسم مقارنة الأرقام (ومدته ٨ دقائق) وقسم مقارنة الأسماء (ومدته ٧ دقائق). ويتطلب الاختبار من الممتحَنين مقارنة عدد من الأزواج التي تتألف من أسماء أوأرقام

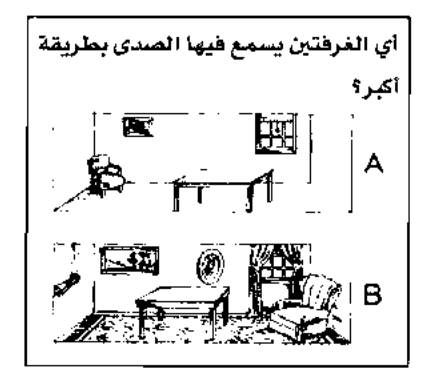
والإشارة إلى الأزواج المتماثلة منها. وبما أن الاختبار يعتمد على السرعة، يجب تصحيح درجات الاختبار من أثر التخمين ثم تحويل الدرجات المصححة إلى معايير مئينية تبعاً للجنس والصف المدرسي (٧-١٢) ولمجموعات المتقدمين إلى الوظائف المكتبية والعاملين فيها. وهنالك ارتباط لا بأس به بين درجات الممتحَنين في اختبار MCT وتقييم المدرسين والمشرفين على أداء الطلاب أوالموظفين في المهام الكتابية. ويساعد برنامج G-7 في إجراء وتصحيح اختبار مشابه للسرعة والصحة.

وبالمقارنة باختبار MCT البسيط، تعد بطارية اختبار القدرات الكتابية واختبار المهام الكتابية العام أدوات قياس أكثر تعقيداً. يتألف اختبار المهام الكتابية العام من ثلاثة اختبارات فرعية تقيم التالي: المهارات الكتابية (التدقيق والترتيب الأبجدي) والمهارات العددية (الحسابات الرياضية وتحديد الأخطاء العددية والبرهان الحسابي) والمهارات اللفظية (التهجئة والقراءة والإستيعاب والمفردات وأصول النحو). وتستغرق الأقسام ٧ دقائق و٢٢ دقيقة ونصف و١٦ دقيقة ونصف على التوالي. ويوفر دليل المستخدم المرفق بالاختبار معلومات عن صدق وثبات الاختبار والمعايير المئينية لعدد من المواقع مقسمة تبعاً لدرجات الاختبارات الفرعية ودرجات الاختبار بأكمله.









الشكل ٦-٣

عينة من أسئلة اختبار بينيه للاستيعاب الميكانيكي. (حقوق الطبع محفوظة لمؤسسة علم النفس المتحدة ١٩٤٢، ١٩٦٧ إلى ١٩٧٠، ١٩٨٠، تمت إعادة الطبع بناءً على موافقة المؤسسة. "BMCT" و"اختبار بينيه للإستيعاب الميكانيكي" وهي علامات مسجلة لمؤسسة علم النفس المتحدة)

مقارنة الأرقام

وضّح أي أزواج الأرقام التالية صحيحاً؟
79524 ______ 79542
5794367 __✓__ 5794367

مقارنة الأسماء

وضّح أي أزواج الأسماء التالية صحيحاً
John C. Linder ______ John C. Lender
Investors Syndicate __✓__ Investors Syndicate
حاول الإجابة على السؤال التالي:
New York World ______ New York World
Cargill Grain Co ______ Cargill Grain Co

الشكل ٦-٤

عينة من أسئلة اختبار مينيسوتا للمهارات الكتابية (حقوق الطبع محفوظة لمؤسسة علم النفس المتحدة عام ١٩٣٣ وجددت عام ١٩٦١، تمت إعادة الطبع بناءً على موافقة المؤسسة)

تتألف بطارية اختبار القدرات الكتابية من سبعة اختبارات فرعية قصيرة هي: الأرشفة ونسخ المعلومات ومقارنة المعلومات واستخدام الجداول وتدقيق النصوص والجمع والطرح والبرهان العددي. مدة الاختبار الكامل هي ستون دقيقة وتتراوح مدد الاختبارات الفرعية بين خمس وعشرين دقيقة. ويضم دليل المستخدم لهذا الاختبار الدرجات المئينية لصورتي الاختبار (أ) و(ب) ولعدد من العينات المختارة من العاملين في الوظائف الكتابية.

القدرات المتعلقة بالحاسب الآلي:

تعتمد معظم الوظائف الكتابية المعاصرة على القدرة على استخدام الحاسب وملحقاته مما يتطلب إدراكاً حسياً سريعاً صحيحاً تدعمه القدرة على التفكير

المنطقي وحل المشكلات. وبناءً على هذا، تم تصميم بطارية اختبار استعداد مستخدمي الحاسب وبطارية اختبار استعداد مبرمجي الحاسب واختبار المهارات CRT بغرض التعرف على الأفراد القادرين على استخدام وبرمجة الحواسب على أكمل وجه. وتتألف بطارية اختبار استعداد مستخدمي الحاسب من ثلاثة اختبارات فرعية تستغرق ٤٥ دقيقة وهي:

(١) تمييز التتابع Sequence Recognition: مدته ١٠ دقائق ويقيم القدرة على التعرف على أنماط التتابع بسرعة.

(٢) تدقيق التصميم Format checking: مدته ٥ دقائق ويقيم القدرة على إدراك تطابق الأرقام والحروف ضمن تصميم معين.

(٣) التفكير المنطقي Logical thinking: مدته ٣٠ دقيقة ويقيم القدرة على تحليل المشاكل وتصور حلول لها.

ويوفر دليل المستخدم المرفق بالاختبار المسالك النموذجية المئوية لدرجات الاختبارات الفرعية ودرجات الاختبار بأكمله التي تم الحصول عليها من عينة صغيرة من مستخدمي الحاسب الخبراء والمتدربين وغير الخبراء.

تتألف بطارية اختبار استعداد مبرمجي الحاسب (CPAB) من خمسة اختبارات فرعية هي:

(١) معاني الألفاظ Verbal meaning: وهي مجموعة المصطلحات المستخدمة في الرياضيات والأعمال وهندسة أنظمة الحواسب (٨ دقائق).

(٢) الاستدلال Reasoning: صياغة الأفكار والتعليمات الكلامية في رموز رياضية (٢٠ دقيقة).

(٣) متسلسلات الحروف Letter series: استخراج الانماط من متسلسلات حروف معطاة (١٠ دقائق).

(٤) القدرة العددية Number ability: المهارة في الحسابات الرياضية (٦ دقائق).

(٥) الرسم التخطيطي Diagramming: حل المشكلات ضمن ترتيب تسلسلي منطقي (٣٥ دقيقة).

تم بناء فقرات هذا الاختبار على عينة من المبرمجين ومحللي النظم الجدد والخبراء. ويوفر دليل المستخدم المرفق بالاختبار المعايير المئينية لدرجات الاختبارات الفرعية ودرجات الاختبار بأكمله لمجموعات من المتدربين وطالبي التوظيف الخبراء والعاملين في الدرجة الأولى من السلم الوظيفي وطلاب كليات المجتمع. كما تتوفر معلومات عن معاملات الصدق التنبؤي للبرمجة في عدد من المؤسسات الإنتاجية والخدماتية. ويسهل توفر صورتين من الاختبار (A و B) عملية إعادة اختبار المتدربين بعد إتمام التدريب.

يُركز اختبار المهارات CRT على القدرة على إدخال المعلومات الأبجدية الرقمية في الحاسب وعلى استدعاء الملفات وتحليلها. ويهدف هذا الاختبار المحوسب إلى التعرف على الأفراد الذين يمكن أن ينجحوا في العمل في مجالات إدخال البيانات وخدمة الزبائن وغيرها من الأعمال المكتبية. كما يقيس القسم الأول من هذا الاختبار سرعة وصحة إدخال البيانات الأبجدية والرقمية بينما يقيس القسم الثاني سرعة وصحة إدخال البيانات الرقمية فقط، ويقيس القسم الثالث القدرة على استدعاء ملفات الزبائن وتحديد الإجابات الصحيحة لعدد من أسئلة الزبائن. وقد تمت دراسات القياس والصدق لهذه الاختبارات المحوسبة على عينة صغيرة نسبياً من الكتّاب وعمال المكاتب الذين يستخدمون الحاسب في أعمالهم اليومية.

بطاريات اختبار الاستعدادات المتعددة

لقد اتضح للعاملين في مجال القياس النفسي والتعليمي أن الإستعانة بمجموعة من الاختبارات (وليس اختبار واحد فقط) ومتغيرات أخرى في عملية

التنبؤ بالأداء الأكاديمي والمهني من شأنه أن يرفع من دقة تنبؤ أدوات القياس، إذ يعتمد النجاح في المدرسة أو العمل على عدد من القدرات المعقدة والعوامل المتداخلة وليس على قدرة أساسية واحدة.

ولا يتوقف النجاح الرياضي أو المهني على الخصائص الإدراكية والنفسية – الحركية للأفراد فقط بل على عوامل أخرى كالشخصية والحافز الداخلي وحتى الحظ والصدف. وعلى حين يمكن التنبؤ بتأثيرات بعض العوامل الآنفة الذكر والسيطرة عليها، تعد بعض هذه العوامل حوادث طارئة خارجة عن سيطرة الفرد. ولهذا السبب لا يمكن لعمليات الإرشاد المهني واختيار الموظفين أن تكون ناجحة تماماً على الرغم من معرفة القائمين بهذه المهام لعناصر النجاح ومحكاته ومتطلباته. وبالرغم من هذا يبقى إجراء بطارية اختبارات قصيرة تقيس قدرات متعددة أفضل بكثير من إجراء اختبار واحد طويل يقيس قدرة واحدة أو قدرتين. وفي حال عدم وجود ارتباط كبير بين نتائج أقسام بطارية الاختبار المختلفة ووجود ارتباط دال فقط بين أقسام الاختبار والمحك المراد التنبؤ به، فإن هذا كافٍ لتقليل الخطأ في التنبؤ. أضف إلى ذلك أن قياس اختبارات البطارية كلها على ذات العينة الممثلة للمجتمع يؤدي إلى الحصول على مجموعة معايير موحدة لجميع أجزاء البطارية، ويمكن بناءً على هذا تقييم درجات الفرد بالمقارنة مع درجات الآخرين وبالمقارنة مع درجاته هو في الأجزاء الأخرى من بطارية الاختبار.

تحليل تباين الدرجات

يمكن إنشاء ملف خاص لدرجات الفرد لبطارية الاختبارات وذلك لمقارنتها مع ملفات درجات الأفراد الذين حققوا نجاحاً في المهن أو البرامج الدراسية المختلفة. ويساعدنا التعرف على الملفات المهنية والأكاديمية التي تشابه ملفات الممتحنين على اختيار الأفراد الذين تتشابه ملفاتهم مع أولئك الذين حققوا نجاحاً في مهنة أو برنامج دراسي معين.

وقبل محاولة تفسير الفروق بين درجات الممتحَن في الاختبارات الفرعية المكونة لبطارية الاستعدادات المتعددة، يجب أن ندرس ثبات الاختبارات والأخطاء المعيارية لقياس تباين درجات كل زوج من الاختبارات الفرعية. وإن ثبات تباين الدرجات (r_{dd}) هو:

$$r_{dd} = (r_{11} + r_{22} - 2r_{12}) / (2 - 2r_{12}) \qquad (١-٦)$$

حيث r_{11} و r_{22} تمثل ثبات الاختبارين و r_{12} هي الارتباط بينهما. ويفرض أنه تم تحويل درجات الاختبارين إلى درجات T قياسية، يكون الخطأ المعياري (Sd) لقياس تباين الدرجات هو:

$$s_d = 10\sqrt{2 - r_{11} - r_{22}} \qquad (٢-٦)$$

ويمكن إجراء هذه الإحصاءات باستخدام برنامج E-7 الموجود على القرص المرن المرافق للكتاب. وإن ثبات التباين بين درجات الاختبارين أقل من ثبات كل من المتغيرين على حدة، والخطأ المعياري لقياس تباين الدرجات أكبر من الخطأ المعياري للقياس لكل اختبار. ويُنصح بأن يساوي تباين درجات أي اختبارين فرعيين من بطارية اختبار خطأ معياري واحد لقياس تباين الدرجات على الأقل أو خطأين معياريين (وهذه أفضل) قبل أن نُخلص إلى القول بأن الفرق ذو دلالة إحصائياً.

مجموعة من اختبارات الاستعدادات المتعددة

تم إعداد بطاريات اختبار الاستعدادات المتعددة أول الأمر لأغراض الإرشاد المهني والتوظيف. ولعل أقدم الاختبارات من هذا النوع بطارية اختبار الاستعداد العام لخدمة التوظيف الأمريكية. وتم في سنوات لاحقة تطوير اختبارات الاستعداد التبايني وغيرها من البطاريات لتأكيد التركيز على الإرشاد الاكاديمي. وتتوفر اليوم في الأسواق أكثر من عشرة بطاريات اختبار للاستعدادات المتعددة

والتي يمكن أن تستخدم في مجالات التوظيف والتعليم على حدٍ سواء. ويحتوي برنامج H-7 على قائمة بأسماء معظم هذه البطاريات (راجع القرص المرن المرفق). وعلى عكس الاختبارات التي تقيس الاستعدادات الواحدة والتي يمكن أن تكون اختبارات أداء أو اختبارات تحريرية، تتميز معظم الاختبارات التي تشكل جزءاً من بطارية اختبار بكونها اختبارات جماعية تُجرى إما عن طريق كراسات الاختبار أو باستخدام الحاسب مما يسهل اختبار أعداد كبيرة من الطلاب أو المتقدمين للوظائف أو المجندين الجدد أو غيرهم في وقتٍ واحد.

اختبارات الاستعداد التبايني

تُستخدم غالبية بطاريات الاختبار التحصيلية في المرحلة الإبتدائية بينما لا تستخدم بطاريات الاستعدادات المتعددة قبل وصول الطالب إلى الصف الثامن أو التاسع. ففي هذه المرحلة العمرية يفترض أن قدرات الطالب الإدراكية قد وصلت إلى درجة من التباين والنضج تسمح لدرجاته في بطارية الاستعداد بأن تساعد في توجيهه مهنياً أو أكاديمياً. إذ تسمح بطاريات الاستعدادات المتعددة للطلاب بالتعرف على نقاط الضعف والقوة لديهم مما يمكن أن يعينهم في إنتقاء المقررات المناسبة لهم أو تحضير أنفسهم لمستقبل مهني معين.

تعد اختبارات الاستعداد التبايني (DAT) أحد أكثر بطاريات الاستعداد الأكاديمية التوجه شعبيةً. وتستخدم هذه البطارية في عمليات التوجيه التربوي والمهني في المراحل المتوسطة والثانوية وفي مقررات التعليم الأساسي للبالغين وكليات المجتمع والمدارس التقنية والمهنية والمؤسسات والبرامج الإصلاحية. تتألف البطارية من ثمانية اختبارات هي:

- اختبار الاستدلال اللفظي (VR).
- اختبار الاستدلال العددي (NR).

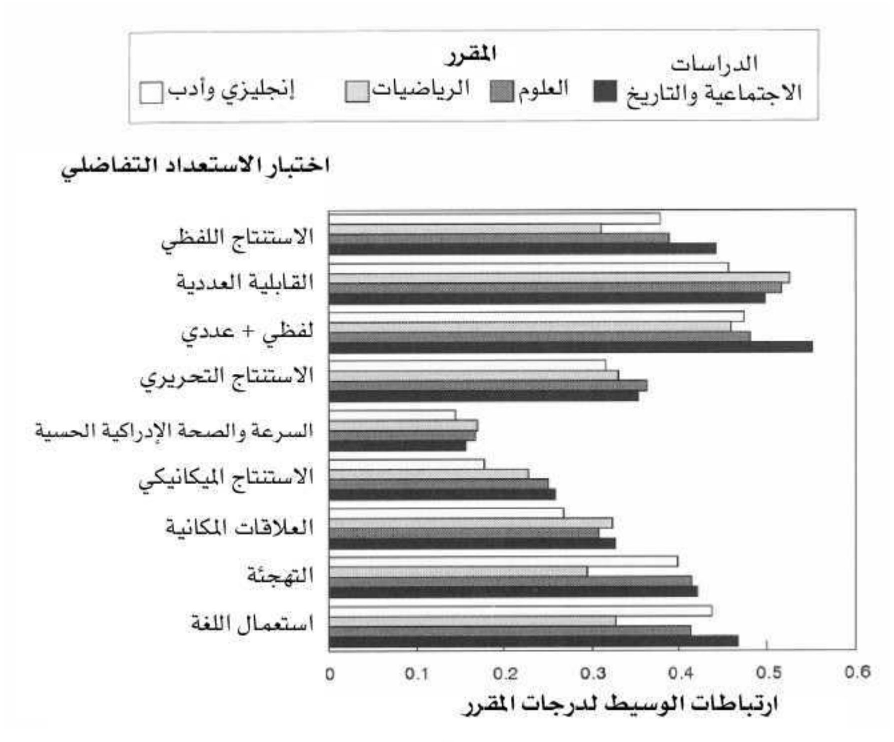
- اختبار الإستدلال التجريدي (AR).
- اختبار السرعة والصحة الإدراكية الحسية (PSA).
- اختبار الإستدلال الميكانيكي (MR).
- اختبار العلاقات المكانية (SR).
- اختبار التهجئة (SP).
- اختبار استعمال اللغة (LU).

وتستغرق أقسام البطارية كلها ساعتين ونصف إلا أن البطارية الجزئية وهي نسخة اختبار تكيفي محوسب من البطارية تستغرق ٩٠ دقيقة فقط. كما تتوفر نسخة من بطارية DAT معدة للاستخدام في مجال الإرشاد المهني حصراً وهي اختبارات الاستعدادات التباينية للتقييم الشخصي والمهني. وتستخدم هذه النسخة من اختبارات DAT مع اختبار قائمة الميول المهنية كجزء من برنامج متكامل لمساعدة الطلاب على إتخاذ قرارات مهنية وتعليمية حكيمة.

وقد تم تقييس الطبعة الخامسة من اختبارات DAT بطريقة جيدة على عينة وطنية من طلاب المدارس المتوسطة والثانوية تم تقسيمها إلى طبقات بناءً على حجم المنطقة التعليمية التي تنتمي إليها المدرسة وعلى الموقع الجغرافي وعلى المستوى الاجتماعي - الاقتصادي وعلى نوع المدرسة (حكوميةوخاصة). ونظراً لارتفاع معدلات الذكور عن معدلات الإناث في اختبارات الإستدلال الميكانيكي والعلاقات المكانية ولإرتفاع معدلات الإناث عن معدلات الذكور في اختبارات السرعة والصحة الإدراكية الحسية واستعمال اللغة، يُعطى مستخدمو اختبارات DAT معايير للجنسين معاً ولكل جنس على حدة.

تترواح معاملات ثبات الإتساق الداخلي للاختبارات الثمانية بين أجزاء ٠,٨٠ و٠,٩٠ بينما تتراوح معاملات ثبات الصور المتكافئة للاختبارات بين

٠,٧٣، و ٠,٩٠، - وتعد نسب الارتباط المثالية بين اختبارات بطارية استعداد هي تلك الأقرب إلى القيمة ٠,٠٠، إلا أن اختبارات DAT ترتبط بنسبة تتراوح بين قريب الصفر و ٠,٧٠ (بين الإستدلال واستعمال اللغة). وتختلف نسب الارتباط بين اختبارات DAT الفرعية ودرجات الطلاب في المقررات الدراسية باختلاف الاختبار الفرعي والمقرر (راجع الشكل ٦-٥). وكما هو متوقع يزداد الارتباط بين الاختبارات الفرعية ذات الاتجاه اللغوي والمقررات اللغوية بينما يزداد الارتباط بين الاختبارات الفرعية ذات الاتجاه الكمي والمقررات العلمية والتقنية. وترتفع نسبة الارتباط بين درجات الطلاب في المقررات المختلفة والاختبارات ذات المحتوى الأكاديمي كاختبارات الاستدلال اللفظي والقدرة العددية والتهجئة واستعمال اللغة. ويعد مركب اختباري الاستدلال اللفظي والقدرة العددية مركباً اختبارياً صادقاً من حيث التنبؤ بدرجات طلاب المدارس الثانوية والجامعات وبنسبة ترابط مرتفعة تتراوح بين ٠,٧٠، و٠,٨٠، وتتميز بعض الاختبارات الفرعية من بطارية DAT بنسب ترابط دالة مع الأداء الوظيفي في عدد من المهن، إلا أن محدودية المعايير تجعل من الأفضل توخي الحذر عند استخدام اختبارات DAT في مجال الانتقاء والتنسيب المهني.



الشكل ٦-٥

ارتباطات الوسيط لاختبارات الاستعداد التبايني العام مع درجات المقررات. (جمعت المعلومات من الطبعة الخامسة من اختبارات الاستعداد التبايني العام. حقوق الطبع محفوظة لمؤسسة علم النفس المتحدة، ١٩٩٠، تمت إعادة الطبع بناءً على موافقة المؤسسة. "DAT" و"اختبارات الاستعداد التبايني العام" هي علامات مسجلة لمؤسسة علم النفس المتحدة)

بطارية اختبار الاستعداد العام

تعود أصول بطارية اختبار الاستعداد العام (GATB) والتي تعد من أكثر بطاريات اختبار الاستعداد الصناعي استخداماً على مر العقود إلى عدد من الاختبارات التي وضعها علماء النفس العاملين في معهد مينيسوتا لأبحاث دعم

التوظيف (MESRI) في الثلاثينيات من القرن الفائت. وكانت بطارية MESRI مؤلفة من عدد من الاختبارات المنفصلة في القدرات العددية والإدراكية الحسية والميكانيكية والنفسية – الحركية، بالإضافة إلى اختبار الذكاء العام. وقد تم إجراء اختبارات البطارية بأكملها على عدد من الكتّاب والميكانيكيين وبائعي المتاجر وغيرهم من أصحاب المهن المختلفة مما أدى بالقائمين على الاختبار إلى التوصل إلى عدد من نماذج القابلية المهنية (OAP) التي يمكن مقارنة أداء كل المتقدمين للاختبار في المستقبل معها. وقد تم استخدام هذه النماذج في قياس بطارية GATB عام ١٩٤٧ من قبل وزارة العمل الأمريكيةبالإستناد إلى تحليل المهن وتحليل العامل لتسعة وخمسين اختباراً.

تم استخدام بطارية GATB بشكل واسع في عمليتي الإرشاد والتنسيب المهني لطلاب الصفوف ٩ إلى ١٢ بالإضافة إلى البالغين. وتعد البطارية جزءاً أساسياً من برنامج الإحالات المهنية لخدمة التوظيف الأمريكية إذ تؤثر درجات الأفراد في الاختبارات على فرص حصولهم على عمل (ديلاهونتي، ١٩٨٨؛ هارتيغان و ويغدور، ١٩٨٩). تتألف البطارية من ١٢ اختباراً، ٨ منها تحريرية و٤ أدائية (تعتمد على الأدوات) تستخدم في أكثر من ٣٥ دولة وتتوفر في ١٢ لغة. تستغرق الاختبارات مجتمعة ساعتين إلا أنه ليس من الضروري إجراء جميع الاختبارات لممتحن ما. وتفيد نتائج الافراد في هذه الاختبارات في تحديد درجات هؤلاء الأفراد في تسعة مجالات استعداد لها أهمية في عالم الأداء المهني وهي: الذكاء والاستعداد اللغوي والاستعداد العددي والاستعداد المكاني وإدراك الأشكال والإدراك الكتابي والتناسق الحركي ودقة الأصابع والدقة اليدوية. وتعد الاختبارات الأدائية الأربعة التي ذكرناها سابقاً والتي تتطلب معدات خاصة ومراقبة من قبل أفراد مدربين من عيوب بطارية GATB وخصوصاً عند مقارنتها مع البطاريات التحريرية كاختبارات الاستعداد التبايني وبطارية القوات المسلحة للاستعداد المهني.

تمت الإستعانة بدرجات عينات من العمال الأمريكيين (العاملين في أكثر من ٨٠٠ مجال عمل مختلف) لقياس بطارية GATB، وتمخض القياس عن استخراج ٣٦ نموذج للقابلية المهنية (OAP) يمكن على أساسها التنبؤ بالنجاح المهني والتعليمي. وعندما يجري الاختبار لشخص ما يتم تحويل درجات هذا الشخص في الاختبارات التسعة إلى رتب مئينية ودرجات قياسية (الوسط الحسابي = ١٠٠، الإنحراف المعياري = ٢٠) ثم تقارن درجات الشخص القياسية مع النماذج الستة والثلاثين للقابلية المهنية.

تم ولسنين طويلة قياس اختبارات GATB على عينات مختلفة عرقياً وذلك كجزء من برنامج العمل الإيجابي Affirmactive action الذي تبنته وزارة العمل الأمريكية والذي يتطلب إعداد معايير مختلفة للفئات العرقية المختلفة(٥). وقد أدى هذا إلى تحديد المعايير المئينية المختلفة للبيض والسود وذوي الأصول الأمريكية اللاتينية المؤلفة من الرتب المئينية للممتحنين بالنسبة إلى فئتهم العرقية. وقد تم بناءً على قانون الحقوق المدنية لعام ١٩٩١ وقف العمل ببرنامج العمل الإيجابي هذا.

تم التوصل عن طريق التحليل العاملي لدرجات ١٢ بطارية اختبارات GATB إلى العوامل الثلاثة التالية:

(١) عامل القدرة العامة على التعلم والذي يتألف من اختبارات لغوية وعددية.

(٢) العامل الإدراكي الحسي والذي يتألف من القدرة المكانية وإدراك الأشكال والإدراك المكتبي.

(٣) العامل النفسي - الحركي والذي يتألف من التنسيق الحركي وبراعة الأصابع والبراعة اليدوية. (هنتر، ١٩٨٣).

---

(٥) Affirmative Action أو Positive Discrimination كما يطلق عليه في المملكة المتحدة هو إجراء يهدف إلى ضمان مشاركة هذه الفئات بنسب مقبولة في مختلف المجالات والمؤسسات والنشاطات. (المترجم).

كما قادت الدراسات المتعددة لثبات بطارية GATB إلى معاملات صور متكافئة تتراوح بين ٠,٨٠، و٠,٩٠، وبقيمة ثبات لاختبارات الحركة أقل من الاختبارات التحريرية. وقد خلُصت عدد من الدراسات التي أجريت لتحديد صدق البطارية إلى أن البطارية هي مقياس تنبؤ فعال بالأداء في عدد من المهن (هارتيغان و ويغدور، ١٩٨٩). إلا أن السرعة التي تتطلبها اختبارات GATB والتي تتحيز للأفراد القادرين على العمل بسرعة تعد من عيوب البطارية التي يضاف إليها كون معايير البطارية قديمة جداً بالإضافة إلى استخدام إجراءات نماذج القابلية المهنية (OAP) متعددة الدرجات الفاصلة في عمليات الإرشاد والتنسيب المهني (أنستازي، ١٩٨٢).

بطارية القوات المسلحة للاستعداد المهني ASVAB:

استخدمت إدارة القوات المسلحة الأمريكية طوال القرن العشرين اختبارات نفسية تهدف إلى قياس القدرة الإدراكية للمجندين، وقد كانت أولى الاختبارات الجماعية التي تم استخدامها هي اختبارات الجيش ألفا وبيتا. صمم اختبار ألفا لقياس القدرات الذهنية العامة للذين يعرفون القراءة والكتابة بينما استخدم اختبار بيتا لقياس قدرات الأميين. وقد تم استبدال اختباري بيتا وألفا خلال الحرب العالمية الثانية باختبارات البحرية والجيش التصنيفية AGCT وNGCT ثم تم تدعيم هذه الاختبارات ومن ثم استبدالها تماماً باختباركفاءات القوات المسلحة (AFQT) وبطارية القوات المسلحة للاستعداد المهني (ASVAB).

ويتم اليوم الاعتماد الكلي على اختبار ASVAB في انتقاء وتصنيف عناصر الجيش والقوات المسلحة في الولايات المتحدة. يتم اجراء اختبار ASVAB بشكل يدوي غير محوسب لأعداد كبيرة من المتقدمين للإلتحاق بصفوف الجيش من البالغين الصغار بالإضافة إلى استخدام هذا الاختبار مع طلاب المدارس الثانوية والسنين الأولى في الجامعة في كافة أنحاء الولايات المتحدة. ويتألف اختبار ASVAB من الاختبارات الفرعية التالية:

١- العلوم العامة (GS): ويتألف من ٢٥ فقرة تقيس معرفة الممتحَن في مجالات العلوم الفيزيائية والطبيعية.

٢- الإستدلال الحسابي (AR): ويتألف من ٣٠ فقرة تقيس قدرة الممتحَن على حل المسائل الرياضية.

٣- معرفة الكلمات (WK): ويتألف من ٣٥ فقرة تقيس قدرة الممتحَن على فهم معاني الكلمات من السياق وعلى التعرف على المرادفات الأنسب لعدد من الكلمات.

٤- استيعاب المقاطع النصية (PC): ويتألف من ١٥ فقرة تقيس قدرة الممتحَن على استخراج المعلومات من النصوص.

٥- العمليات العددية (NO): ويتألف من ٥٠ فقرة تقيس قدرة الممتحَن على القيام بالعمليات الحسابية.

٦- سرعة الترميز (CS): ويتألف من ٨٤ فقرة تقيس قدرة الممتحَن على تحديد رموز عددية للكلمات باستخدام مفتاح للرموز.

٧- المعلومات الميكانيكية (AS): ويتألف من ٢٥ فقرة تقيس قدرة الممتحَن على التعامل مع الأدوات الميكانيكية والسيارات والتعرف على مصطلحات المهنة وممارساتها.

٨- المعرفة الرياضية (MK): ويتألف من ٢٥ فقرة تقيس معرفة الممتحَن للمبادئ الأساسية في الرياضيات[1].

---

(١) على ما يبدو أن هناك خطأ مطبعي في النسخة الإنجليزية (المترجم).

٩- الاستيعاب الميكانيكي (MC): ويتألف من ٢٥ فقرة تقيس معرفة الممتحن للمبادئ الفيزيائية والميكانيكية وقدرته على تصور كيفية عمل بعض الأدوات التي يضم الاختبار صوراً لها.

١٠- المعلومات الألكترونية (EI): ويتألف من ٢٥ فقرة تقيس معرفة الممتحن للمبادئ الكهربائية والألكترونية.

وكما هو موضح بهذه التوصيفات وعينات الأسئلة في الشكل ٦-١ فإن معظم اختبارات ASVAB تقيس في الواقع تحصيل الطلاب وخبراتهم الدراسية. وتتراوح أوقات الاختبارات بين ثلاث دقائق لاختبار العمليات العددية وسبع دقائق لاختبار سرعة الترميز (وهما اختبارا السرعة في البطارية) إلى ست وثلاثين دقيقة لاختبار الإستدلال الحسابي. وتستغرق البطارية بأكملها ١٤٤ دقيقة.

يتم تحويل الدرجات الأولية لاختبارات البطارية العشرة وثلاثة اختبارات أكاديمية مركبة (أي تضم عدد من اختبارات البطارية التي تقيس المهارات الأكاديمية وليس المهنية) إلى درجات Tscores (الوسط = ٥٠ والإنحراف المعياري = ١٠) وإلى رتب مئوية. ويتم حساب نقاط الدرجات الأكاديمية المركبة على النحو التالي:

القدرة اللغوية = معرفة الكلمات + استيعاب المقاطع النصية

القدرة الرياضية = الإستدلال الحسابي + المعرفة الرياضية

القدرة الأكاديمية = القدرة اللغوية + العمليات العددية

كما يمكن أيضاً حساب درجات اختبارات مركبة لمجموعة مركبة من أربعة مهن وذلك بتركيب مناسب لنقاط درجات عشرة اختبارات أساسية تضم: اختبار الميكانيك والصناعات اليدوية، واختبار الأعمال والوظائف الكتابية، واختبارات المعرفة الألكترونية والكهربائية، واختبار الصحة والمجتمع والتقنية.

ويوضح الشكل ٦-٧ تمثيلاً بيانياً لدرجات على بطارية القوات المسلحة للاستعداد المهني ASVB لطالب ثانوي وهمي تم التعبير عنها من خلال حزم درجات مئينية تدل على المجالات التي يحتمل أن تقع ضمنها درجات الطالب الحقيقية. كما يعطي الشكل تمثيلاً لدرجات T لطلاب من فصل واحد/ ومن جنس واحد ومن فصل واحد/ ومن جنسين، هذا بالإضافة إلى حزم درجات الطلاب المئينية من فصل واحد وجنس واحد. كما يضم ملف الطالب لاختبارات ASVAB رمزيّ الطالب للمرحلتين الإبتدائية والثانوية ودرجاته في اختبار المهن العسكرية. ويتم تفسير الرموز بالإستعانة بأداة OCCU-FIND الموجودة في الدليل العملي المصمم لمساعدة الممتحَنين في التعرف على المهن العسكرية الأكثر مناسبة لهم.كما تساعد درجات الممتحَن في اختبار المهن العسكرية والذي يستخدم مع الرسومات البيانية في كتيب المهن العسكرية في تقدير فرص الممتحَن في النجاح في بعض المهن العسكرية التطوعية.

تتمتع اختبارات ASVAB بدرجة عالية من الثبات وتتراوح نسب ثبات الإتساق الداخلي في هذه الاختبارات بين ٠.٩٢ و٠.٩٦ بينما تتراوح نسبة ثبات الصور الاختبارية المتكافئة بين ٠,٧٧ و٠,٨٥ للاختبارات التي تتطلب سرعة (العمليات العددية وسرعة الترميز) و٠.٧١ و٠.٩١ للاختبارات الباقية. وتتوفر معلومات كاملة وشاملة عن صدق اختبارات ASVAB عند استخدامها مع العاملين في قطاع الجيش والقوات المسلحة ومع طلاب المدارس الثانوية والكليات (وزارة الدفاع الأمريكية، ١٩٩٣، ١٩٩٤). ولاختبارات ASVAB عدد من العيوب من أهمها أن تكرار الاختبار ذاته في عدد من مركبات الاختبار يعني ارتفاع نسبة الترابط بين هذه المركبات مما يجعل بعض النتائج التي توفرها الاختبارات عن الطالب نتائج مكررة تشير إلى القدرة الإدراكية العامة للطالب وليس إلى عدد من الاستعدادات الخاصة. ومن العيوب الأخرى لهذه الاختبارات كونها تحمل كمية كبيرة من عامل القدرة العام[٦] general ability factor(g) مما أدى بعدد من الخبراء إلى الحكم

(٦) أي إنها تقييم القدرات العامة وليس القدرات الخاصة المتعددة للأفراد. (المترجم).

على اختبارات ASVAB بأنها فشلت في إعطاء معلومات تمييزية عن المهارات الفردية المختلفة مما يجعل هذه الاختبارات غير مفيدة في التنسيب الناجح للأفراد في الوظائف المختلفة (راجع مورفي، ١٩٨٤؛ وثورندايك، ١٩٨٥).

وقد تم اقتراح عدد من التوصيات لرفع مستوى فعالية اختبارات ASVAB في التنبؤ بالأداء المهني والتنسيب (ويلش، كوتشينكاس وكورران، ١٩٩٠؛ زايدنر وجونسون، ١٩٩٢) إلا أن نجاح هذه التوصيات لا يزال أمراً غير معروف. ومن الجدير بالذكر أن اختبار ASVAB-CAT هو أحد النسخ الجديدة من بطارية ASVAB والتي تعتمد على طريقة الاختبارات التكيفية المحوسبة والتي تتميزبكونها تتطلب وقتاً أقصر وتعطي نتائج أكثر دقة في المستويات الأعلى والأدنى للأداء وتغذيتها الراجعة فورية وأوقات إجرائها مرنة (لارسون، ١٩٩٤). ولا يمكن الجزم هنا بمدى كون هذه المزايا تبرر الثمن الباهظ للاختبارات التكيفية المحوسبة.

بطاريات الاستعدادات المتعددة:

تضم العديد من الاختبارات النفسية كلمة "استعداد" في أسمائها مثل: مسح استعدادات الموظفين، واختبار فلانيغان لتصنيف الاستعداد، وبطارية الاستعداد المتعدد الأبعاد، واختبار الاستعداد المدرسي. وإن أقدم هذه الاختبارات وأطولها هو اختبار فلانيغان لتصنيف الاستعداد (FACT) والذي يتألف من ١٦ اختباراً مختلفاً تستغرق بين دقيقتين إلى أربعين دقيقة. ويقيس الاختبار استعدادات العاملين في الوظائف الصناعية والميكانيكية ذات المستوى المبتدئ من حيث: الحساب والتجميع والترميز وقطع الآلات والتناسق الحركي والتعبير والإبداع والتحقق والكشف والحكم والإستيعاب والميكانيك والذاكرة والأنماط والدقة والإستدلال والتعامل مع المقاييس والجداول. وكما يظهر من الأسماء، تركز هذه الاختبارات على المهارات الصناعية. وتعد بطارية أخرى من اختبارات فلانيغان وهي اختبارات فلانيغان الصناعية (FT) بطارية تقيس المهارات الأكثر تطوراً من

تلك التي تقيسها اختبارات (FACT) بالإضافة إلى أنها تطالب الممتحن بإنهاء الاختبار في وقت أقصر. وتعد ١٥ من اختبارات FIT الثمانية عشرة نسخ معدلة عن اختبارات FACT إذ تم تعديل اختبار الإستدلال إلى اختبار في الإستدلال الرياضي وأضيف اختباران للمفردات والمعرفة الإلكترونية. وتدخل بطاريتا الاختبار FACT وFIT في إطار برنامج للبحث في مجال الاختبارات اعتبر طموحاً في ذلك الوقت حيث بلغ أوجه في الخمسينيات والستينيات إلا أن العمل فيه لم يكتمل. ولذلك يبقى صدق هذه البطاريات ضعيفاً ومعاييرها قديمة وبالية.

الشكل ٦-٦

عينة لفقرات من بطارية القوات المسلحة للاستعداد المهني. (تمت إعادة الطبع عن دليل المرشد ١٩/١٨ ASVAB بموافقة وزارة الدفاع الأمريكية)

العلوم العامة

١- يلقي كسوف الشمس بظل

١- أ) القمر على الشمس

١- ب) القمر على الأرض

١- ج) الأرض على الشمس

١- د) الأرض على القمر

٢- يطلق على العناصر التي تسرِّع التفاعلات الكيميائية دون أن تتأثر هي نفسها بالتفاعلات اسم:

٢- أ) المصدات

٢- ب) الجزيئات

٢- ج) المثبطات

٢- د) المحفزات

الاستدلال الحسابي

٣- كم من الباصات التي تتسع لـ ٣٦ راكباً نحتاج لنقل ١٤٤ شخصاً؟

٣- أ) ٣

٣- ب) ٤

٣- ج) ٥

٣- د) ٦

٤- تكلف معالجة ياردة مربعة من القماش الخام لتصبح مقاومة للماء ٠,٥٠ دولار. ما كلفة معالجة غطاء لسيارة قياساته ١٥ و٢٤ إنشاً مقدرة بالدولار.

٤- أ) ٦,٦٧

٤- ب) ١٨,٠٠

٤- ج) ٢٠,٠٠

٤- د) ١٨٠,٠٠

معرفة الكلمات

٥- الريح اليوم متغيرة، أي:

٥- أ) معتدلة

٥- ب) ثابتة

٥- ج) متبدلة

٥- د) باردة

٦- الأساسيات Rudiments هي كلمة تعني تقريباً:

٦- أ) السياسة

٦- ب) التفاصيل

٦- ج) فرص الترقية

٦- د) الإجراءات والطرق الأساسية

استيعاب المقاطع النصية

٧- تعزى ٢٥٪ من السرقات إلى الشبابيك أو الأبواب غير المقفلة. إن الجريمة هي فرصة ورغبة. ولمكافحة الجريمة، يتوجب على الأفراد:

٧- أ) توفير الرغبة

٧- ب) توفير الفرصة

٧- ج) تفويت الرغبة

٧- د) تفويت الفرصة

٨- تندر المياه في عدد كبير من مناطق العالم لدرجة تبذل معها كل المحاولات لعدم هدرها. ففي أحدى الواحات في الصحراء الكبرى تم تحديد كمية الماء التي تحتاجها كل نخلة. ما كمية الماء التي يجب توفيرها لكل شجرة؟

٨- أ) لا ماء

٨- ب) الكمية اللازمة

٨- ج) ماء كل يومين

٨- د) ماء للأشجار السليمة فقط

العمليات العددية

٩- ٣ + ٩ =

٩- أ) ٣

٩- ب) ٦

٩- ج) ١٢

٩- د) ١٢

١٠- ٦٠ ÷ ١٠=

١٠- أ) ٤

١٠- ب) ٤

١٠- ج) ٨

١٠- د) ١٠

سرعة الترميز

المفتاح

صفقة = ٨٣٨٥	لعبة= ٦٤٥٦	سكينة= ٧١٥٠	كنبة=٩٦٤٥
موسيقا= ١١١٧	ذقن= ٨٩٣٠	بيت= ٢٨٥٩	مالك= ٦٢٢٧

الأسئلة	الأجوبة				
	أ	ب	ج	د	هـ
١١- لعبة	٦٤٥٦	٧١٥٠	٨٣٨٥	٨٩٣٠	٩٦٤٥
١٢- سكينة	١١١٧	٦٤٥٦	٧١٥٠	٦٢٢٧	٨٣٨٥
١٣- صفقة	٢٨٥٩	٦٢٢٧	١١١٧	٨٣٨٥	٩٦٤٥
١٤- ذقن	٢٨٥٩	٦٤٥٦	٨٣٨٥	٨٩٣٠	٩٦٤٥
١٥- بيت	١١١٧	٢٨٥٩	٦٢٢٧	٧١٥٠	٦٤٥٦
١٦- كنبة	٧١٥٠	١١١٧	٨٣٨٥	٨٩٣٠	٩٦٤٥
١٧- مالك	٢٨٥٩	٦٢٢٧	٦٤٥٦	٧١٥٠	٨٩٣٠

معلومات ميكانيكية

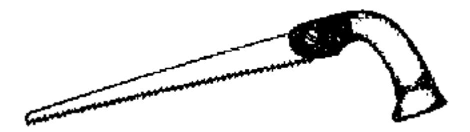
١٨- تستهلك السيارة زيت أكثر عندما يبلى:

١٨- أ) البستونات

١٨- ب) حلقات البستونات

١٨- ج) نواقل الحركة

١٨- د) قضبان الوصل



١٩- يستخدم المنشار الموجود في الصورة لقطع:

١٩- أ) الخشب الرقائقي

١٩- ب) الثقوب الغريبة الشكل في الخشب

١٩- ج) الخشب على مسار حبيبات

١٩- د) الخشب من خلال الحبيبات

المعرفة الرياضية

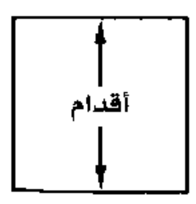
٢٠- إذا كانت س ÷ ٦=٧، إذاً س تساوي:

٢٠- أ) -١

٢٠- ب) ٠

٢٠- ج) ١

٢٠- د) ٦/٧



٢١- ما مساحة المربع؟

٢١- أ) ١ قدم مربعة واحدة

٢١- ب) ٥ أقدام مربعة

٢١- ج) ١٠ أقدام مربع

٢١- د) ٢٥ قدماً مربعة

الاستيعاب الميكانيكي

٢٢- أي من العمودين يحمل الوزن الأكبر من الحمولة؟

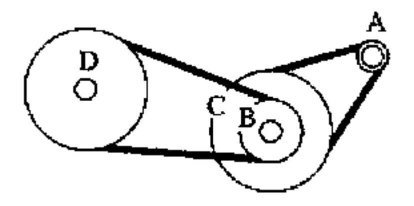


٢٢- أ) العمود A

٢٢- ب) العمود B

٢٢- ج) الوزن متساوي على كلا العمودين

٢٢- د) لا يمكن تحديد الإجابة



٢٣- أي من أقراص الشد أعلاه تدور بأكبر سرعة؟

٢٣- أ) A

٢٣- ب) B

٢٣- ج) C

٢٣- د) D

المعلومات الكهربائية

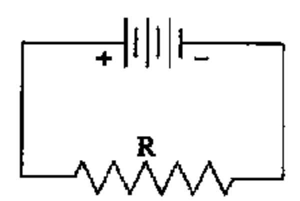
٢٤- أي من المواد التالية الأضعف مقاومة؟

٢٤- أ) الخشب

٢٤- ب) الحديد

٢٤- ج) المطاط

٢٤- د) الفضة



٢٥- في الدارة أعلاه المقاومة تساوي ١٠٠ أوم والتيار يساوي ٠,١ أمبير. الفولتية تساوي:

٢٥- أ) ٥ فولط

٢٥- ب) ١٠ فولط

٢٥- ج) ١٠٠ فولط

٢٥- د) ١٠٠٠ فولط

SMITH KRIS
GRADE 11
SSN 987-65-4321
TEST DATE 9/15/95
LINCOLN HIGH SCHOOL
LINCOLN

SEX F
IL

ASVAB
ARMED SERVICES VOCATIONAL APTITUDE BATTERY

Percentile Scores

ASVAB Scores	Same Grade/Same Sex	Same Grade/Same Sex Percentile Score Bands	Same Grade/Opposite Sex
		1 5 10 20 30 40 50 60 70 80 90 95 99	
Academic Ability (AA = VA + MA)	75		73
Verbal Ability (VA = WK + PC)	71		73
Math Ability (MA = AR + MK)	78		76
Word Knowledge (WK)	75		74
Paragraph Comprehension (PC)	67		74
Arithmetic Reasoning (AR)	74		66
Mathematics Knowledge (MK)	78		82
General Science (GS)	60		47
Auto & Shop Information (AS)	44		21
Mechanical Comprehension (MC)	55		31
Electronics Information (EI)	73		49
Numerical Operations (NO)	74		87
Coding Speed (CS)	44		70
ASVAB Codes 2,3 Military Careers Score 203		1 5 10 20 30 40 50 60 70 80 90 95 99	
(SEE BACK OF SHEET FOR EXPLANATION)			

الشكل ٦,٧

مخطط بياني للدرجات من بطارية القوات المسلحة للاستعداد المهني. (تمت إعادة الطبع بموافقة وزارة الدفاع الأمريكية من دليل المستخدم للمرشد المهني ASVAB 18/19).

اختبار استعداد الموظفين

يشابه محتوى اختبار استعداد الموظفين (EAS) والذي نشر في الثمانينيات محتوى اختبارات الاستعداد التباينية. ويساعد اختبار EAS في اختيار الموظفين المناسبين للعمل في مجالات البيع والأعمال الكتابية والصناعية، ويتألف الاختبار من عشر اختبارات فرعية هي: الإستيعاب اللغوي والقدرة العددية والملاحقة البصرية والسرعة البصرية والصحة وتخيل الفراغ والإستدلال العددي والإستدلال اللفظي والطلاقة الكلامية والسرعة اليدوية والإتقان والإستدلال الرمزي. وتعتبر الاختبارات العشرة اختبارات سرعة إذ لا يستغرق واحدها أكثر من ٥ دقائق. وكما

يلاحظ من أسماء الاختبارات الفرعية المدرجة أعلاه، تركز بطارية EASعلى عناصر الإستدلال اللفظي والعددي مما يجعلها أداة تقييم مفيدة ليس في إطار الإرشاد والإنتقاء المهني فحسب بل في الإرشاد الأكاديمي كذلك.

اختبار الاستعدادات المتعددة الأبعاد

يمكن اعتبار بطارية الاستعدادات المتعددة الأبعاد (MAB) نسخة معدلة، تحريرية وجماعية عن اختبار قياسي آخر هو مقياس ويكسلر المعدل لذكاء البالغين (WAIS-R) (راجع الفصل السابع من الكتاب). يتألف اختبار MAB من خمسة اختبارات فرعية لفظية (المعلومات والإستيعاب والحساب والمتشابهات والمفردات) وخمسة اختبارات فرعية أدائية (تكميل الصور،ترتيب الصور،تصميم المكعبات، تجميع الأشياء،الترميز،المتاهات) مدة كل منها ٧ دقائق. وتُجمع درجات الاختبارات الخمسة الأولى لتمثل معدل الذكاء اللفظي للممتحن بينما تُجمع درجات الاختبارات الخمسة الباقية لتمثل معدل ذكائه الأدائي، وتُجمع الدرجتان للحصول على المقياس الكامل لذكاء الممتحَن. أما معاملات ثبات الاختبار وإعادة الاختبار لمعدلات الذكاء اللفظية والأدائية ومعدل المقياس الكامل فتزيد، بحسب تقارير المسؤولين عن الاختبار، عن ٠,٩٠ كما بينت الدراسات وجود ارتباط عالٍ بين الأداء على اختبار MAB والأداء على مقاييس الذكاء الأخرى.

اختبار التقدير المدرسي

عرف اختبار SAT وحتى السنوات القليلة الأخيرة باسم اختبار الاستعداد المدرسي والذي خضع إليه عدد هائل من طلاب المدارس الأمريكية الراغبين في الإلتحاق بالجامعات ومنذ العشرينيات من القرن الفائت. إلا أنه، وبسبب سوء الفهم والنقد اللاذع لما يقيسه اختبار SAT تم تعديل الاختبار وتغيير اسمه إلى "اختبار التقدير المدرسي" والمحافظة على الاسم المختصر SAT. ولا يختلف

اختبار SAT عن غيره من اختبارات الأداء (كاختبار سجلات الخريجين GRE واختبار الاستعداد الإدراي للخريجين GMAT واختبار القبول في كلية الحقوق LSAT واختبار القبول في كلية الطب MCAT)، في كون نتائجه تستخدم كأدوات تنبؤ بالأداء. ويمكن للقارئ الحصول على قائمة وصفية لهذه الاختبارات وغيرها من اختبارات القبول الجامعي والتخصصي عن طريق الضغط على الخيار الثاني من القائمة في برنامج H-7.

وتتألف النسخة الحالية من اختبار SAT من قسمين: SAT I للإستدلال، وSAT II لاختبارات المواد المختلفة. يتألف القسم الأول من جزئين يقيسان القدرة على الاستدلال الرياضي واللفظي ويستغرق كل واحد منهما ٧٥ دقيقة. يتألف قسم الاستدلال اللفظي من ٧٨ سؤال من نوع الخيارات المتعددة عن المتشابهات وإكمال الجمل والقراءة النقدية، بينما يتألف قسم الإستدلال الرياضي من ٦٠ سؤال في الرياضيات العامة والمقارنات الكميّة والإجابات التي يتوصل إليها الطلاب أنفسهم. ويُسمح للطلاب باستخدام الآلات الحاسبة في النسخة الحالية من اختبار SAT وهذا أمر لم يمكن مسموحاً في النسخ السابقة. ويتم تحويل الدرجات الخام في أقسام SAT I الرياضية واللغوية إلى درجات قياسية على مقياس CEEB (بوسط حسابي = ٥٠٠ وإنحراف معياري = ١٠٠ ومدى (مجال) يمتد بين ٢٠٠ إلى ٨٠٠).

تعتبر الاختبارات اللغوية والرياضية من النسخة المعدلة من اختبار SAT أقرب إلى الشق (الطرف) الاستعدادي للمجال الاستعدادي - التحصيلي، بينما تحتل اختبارات المواد الدراسية المختلفة لاختبار SAT II مواضع أقرب للشق التحصيلي من المجال ذاته. وتعتمد الاختبارات المنتقاة من اختبار SAT II على رغبة الطالب الذي يمكن أن يطلب أن يتم اختباره في اللغة الأجنبية أو التاريخ أو الرياضيات أو غيرها من المواد الأكاديمية. وأخيراً؛ يتم أخذ عينة من كتابات

الطالب، وتشكل أسئلة الاختيارات في مجالات اللغة المكتوبة والبلاغة والتعبير المنطقي جزءاً من اختبار SAT II، ويتم وضع درجات اختبار SAT IIعلى مقياس CEEB شأنه كشأن اختبار SAT I.

وعلى الرغم من تغيير اسم اختبار SAT وتعديل محتواه وطريقة عمله، لا يزال العديد من الناس مستاؤون من النفوذ الكبير والأثر العميق للاختبارات النفسية التقليدية والحديثة على حياة الطلاب والموظفين ومستقبلهم، وقد قاد هذا في عديد من الحالات إلى دعاوى قضائية ونزاعات بالإضافة إلى الإنتقادات التي توجه في وسائل الإعلام إلى الاختبارات التعليمية والنفسية. ومن المفيد للعاملين في مجال هذه الاختبارات محاولة الإستفادة من النقد الموجه إليهم وذلك في إعادة النظرفي مفاهيم وطرق وحتى أنواع الاختبارات المستعملة في تقييم السلوك البشري والإدراك ومدى جدوى نتائج هذه الاختبارات.

صدق اختبارات الاستعداد

ترتبط فعالية اختبارات الاستعداد كأداة للاختيار والتسيب والتصنيف المهني وكأداة مساعدة في عمليات الإرشاد والتخطيط التربوي والمهني بقدرة هذه الاختبارات على التنبؤ بشكل دقيق بأداء الأفراد ضمن حيز الواقع العملي. ويعد الصدق التنبؤي للاختبارات والذي تحدثنا عنه في الفصل الثاني في هذا الكتاب والفصلين الرابع والخامس من كتاب سلالم الرصد وقوائم التقدير(أيكين، ١٩٩٦) أمراً هاماً جداً في اختبارات الاستعداد. وبالرغم من أن الهدف الأساسي لاختبارات الاستعداد هو التنبؤ بأداء الأفراد في المدرسة والوظيفة وغيرها من أجزاء الحياة العملية، فإن الجدول ٦-١ يوضح أن درجة الإرتباط بين اختبارات الاستعداد المتعددة ومحكات التدريب أو الكفاءة في العمل ليست عالية. ويبقى مدى الصدق بين درجات اختبارات الاستعداد للتنبؤ بمدى التحصيل في مجالات التدريب والتعليم أعلى من تلك المتعلقة بالأداء الوظيفي، إلا أنه، وفي الحالتين،

لا تعد درجة الصدق هذه عالية. وبالنظر إلى معدل معاملات الصدق لعدد من اختبارات الاستعداد التي تتنبأ بالأداء في عدد من المجالات المهنية لوحظ أن هذه المعاملات غالباً ما تقارب ٠,٢٠ ونادراً ما تتجاوز ٠,٣٠ (جيسيللي، ١٩٧٣). وتتغيرنسبة المعاملات بتغير المحك (مثال: درجات الطلاب في برنامج تدريبي أو تقييمات الأداء الوظيفي) والموقف وعينة الاختبار والفترة الزمنية بين إجراء الاختبار وجمع معلومات المحك. كما أفادت نتائج التحقيقات المسحية (كولي ولونز، ١٩٦٨؛ ثورندايك وهاغن، ١٩٥٩) أن درجات اختبارات الاستعداد غير صحيحة في التنبؤ بالخيارات الوظيفية أو بمستويات الأداء في حرفة معينة. ويفيد أخذ الاهتمامات والدوافع وغيرها من المتغيرات الفردية ومتغيرات المواقف بعين الاعتبار في تحسين التنبؤ بالاختيارات الوظيفية والنجاح في العمل ورضا الأفراد على اختيارتهم المهنية. وفي هذه الحالة لا يتم الاعتماد على درجات اختبار الاستعداد فقط مما يرفع درجة الارتباط المتعددة الناتجة إلا أن هذه النسبة نادراً ما تتجاوز ٠,٥٠ كما يمكن للمتغيرات غير الإدراكية كالاهتمامات الفردية أن تقوم بعمل المتغير الوسيط في الارتباط بين درجات اختبار الاستعداد ومقاييس الإنتاجية، فعلى سبيل المثال يمكن للارتباط بين درجات الاستعداد والإنتاجية أن تكون أدنى مع مجموعة من الأفراد على درجة عالية من الاهتمام بالعمل منها مع مجموعة على درجة أعلى من الاهتمام.

وتلعب عدة عوامل غير الخصائص السايكومترية للاختبار دوراً في صعوبة التنبؤ بالأداء الوظيفي بالإستناد إلى اختبارات الاستعداد. من هذه العوامل: مشكلات تحديد وقياس معايير النجاح الوظيفي والأمور الطارئة وغير المتوقعة التي قد تؤثر على المؤسسات والافراد (كالتغيرات الإقتصادية أو الظروف الاجتماعية) بالإضافة إلى عوامل يلعب الحظ فيها دوراً كبيراً كالزواج من ابنة رب العمل أو ربح اليانصيب أو التعرض إلى إصابة بليغة أو مرض مزمن.

الجدول ٦-١
مدى الارتباط بين اختبارات القدرة ومحكات المهنة

	المقدرة							
فئات المهنة	تفكيري		مكاني وميكانيكي		صحة إدراكية		حركي	
	أ	ب	أ	ب	أ	ب	أ	ب
كتَبه	٠,٤٧	٠,٢٨	٠,٣٤	٠,١٧	٠,٤٠	٠,٢٩	٠,١٤	٠,١٦
مديرون	٠,٣٠	٠,٢٧	٠,٢٨	٠,٢٢	٠,٢٣	٠,٢٥	٠,٠٢	٠,١٤
حماية	٠,٤٢	٠,٢٢	٠,٣٥	٠,١٨	٠,٣٠	٠,٢١		٠,١٤
مبيعات		٠,١٩		٠,١٨		٠,٠٤٠		٠,١٢
خدمات	٠,٤٢	٠,٢٧	٠,٣١	٠,١٣	٠,٢٥	٠,١٠	٠,٢١	٠,١٥
تجارة وحرف	٠,٤١	٠,٢٥	٠,٤١	٠,٢٣	٠,٣٥	٠,٢٤	٠,٢٠	٠,١٩
مشغلو مركبات	٠,١٨	٠,١٦	٠,٣١	٠,٢٠	٠,٠٩٠	٠,١٧	٠,٣١	٠,٢٥

ملاحظة:

أ = محكات تدريبية؛ ب = محكات كفاءة

المصدر: بيانات من جيسلي Ghiselli، ١٩٧٣

نماذج التنبؤ

على الرغم من أن اختبارات الذكاء أو اختبارات القدرة العامة أقدر على التنبؤ بأداء الأفراد التعليمي والمهني من العديد من الاختبارات التي تقيس الاستعدادات الخاصة (كارول، ١٩٧٨؛ ماكنيمار، ١٩٦٤)، تُمكننا عملية اختيار حكيمة لاختبارات الاستعدادات الخاصة يتبعها استخدام إجراءات إحصائية معقدة للجمع بين درجات هذه الاختبارات من تقليل نسبة الخطأ في التنبؤ. وقد تم الاعتماد على استراتيجيتين للاختيار عند استخدام عدد من اختبارات الاستعداد وغيرها من أدوات التنبؤ بالأداء: (١) استراتيجية الحدود (الدرجات) الفاصلة المتعددة Multiple - cutoff والتي تتطلب من الممتحنين الحصول على حد أدنى من الدرجات في كل من الاختبارات وأدوات القياس التي يخضع لها قبل أن يتم قبوله في المدرسة أو الجامعة أو الوظيفة؛ و(٢) استراتيجية الانحدار المتعدد multiple - regression والتي تستخدم تركيبة من المتغيرات الخطية المرجحة تم تصميمها للحصول على أقل نسبة ممكنة من خطأ التنبؤ. ولا تخلو الاستراتيجيتان أعلاه من العيوب إذ تم انتقاد طريقة الدرجات الفاصلة المتعددة لكونها لا تسمح للدرجات العالية في أحد الاختبارات أن تعوض عن النقص في درجات اختبار آخر، كما تم انتقاد طريقة الإنحدار المتعدد لافتراضها بأن متغيرات للتنبؤ مرتبطة خطياً (إما كل متغير على حدة أو مجتمعاً مع غيره من المتغيرات) مع محكات الأداء. كما تم استخدام أنواع أخرى من النماذج الإحصائية المعقدة كالتحليل التمييزي والتحليل القانوني وتحليل المسار ونمذجة المعادلات البنائية في عملية التنبؤ إلا أن هذه الاستخدامات تبقى محدودة إذا ما قورنت باستخدامات طريقتي الحدود (الدرجات) الفاصلة المتعددة و طريقة الإنحدار المتعدد.

وهنالك حالات لا يمكن فيها استخدام اختبار ما على الرغم من وجود إرتباط وثيق بين درجات الاختبار ومقاييس الأداء لمحك ما، إذ من الممكن جداً أن

يكون هذا الاختبار لا يزيد شيئاً على إجراءات الاختيار المستخدمة أصلاً من حيث التنبؤ بأداء الأفراد في محك ما؛ يعني هذا أن الاختبار لا يتميز بالصدق المتزايد (incremental validity). ومن المهم عند اختيار اختبارات الاستعداد مراعاة كون هذه الاختبارات قادرة على التنبؤ ولوحدها بأداء الأفراد في الدورات التدريبية أو في الوظائف المختلفة. كما من الممكن جداً أن يتميز اختبار ما بالصدق المتزايد المذكور أعلاه إلا أن كلفة هذا الاختبار قد تكون باهظة جداً مما لا يبرر استخدامه. فمما لا شك فيه أن الربح المادي هو الهدف الرئيسي لشتى أنواع الشركات والمؤسسات ولا يمكن لأي من هذه الشركات أن تحافظ على بقائها إذا تجاوزت نفقاتها الأرباح التي تحققها. وختاماً تجدر الإشارة إلى أنه تتوفر اختبارات صادقة ومعقولة التكاليف إلا أنها متحيزة وغير عادلة.

عدل الاختبارات

يشترط في الاختبارات التي تستخدم في الاختيار بين مجموع المتقدمين أن تكون منصفة لجميع الفئات التي يتم اختبارها. ويسبب تنوع الخلفيات الاجتماعية والحضارية للممتحَنين وتنوع التجارب التي يعيشونها تفاوتاً في الدرجات التي يمكن أن يحصل عليها عرق أو جنس ما في اختبارات الاستعداد المختلفة. وبينما لا يعد هذا مؤشراً على تحيز الاختبار، يدل ارتفاع معدل الارتباط بين الاختبار ودرجات المحك عند مجموعة بشرية معينة عن غيرها من المجموعات على وجود تحيز في الاختبار. ويعتبر التعريف الأكثر شيوعاً واستخداماً لعدل اختبارات الاستعداد هو التعريف الذي يقول بأن اختبار الاستعداد يكون عادلاً فقط عندما الذي يتنبأ بالأداء المحكي لجميع الفئات التي يتم اختبارها بنفس الدرجة من الدقة. وقد أدى الشرط القانوني بكون الاختبارات التي تستخدم في اختيار الموظفين والطلاب اختبارات عادلة إلى العديد من الدعاوى والنزاعات القانونية والقضائية في السنوات الأخيرة.

واستناداً إلى توجيهات الموحدة لإجراءات اختيار الموظفين والتي وضعتها لجنة فرص العمل المتكافئة (EEOC) يتحتم على عملية إجراءات اختيار الموظفين المشروعة قانونياً أن تقيس ما لا يقل عن ٥٠٪ من المهارات والقدرات التي يتطلبها العمل. كما تشترط التوصيات استخدام قاعدة الأربعة أخماس والتي تضمن عدم اختيار أفراد ينتمون إلى مجموعة بشرية معينة بنسب عالية جداً عن غيرها من المجموعات وهو ما يطلق عليه اسم "التأثير السلبي" للاختبارات. وتنص قاعدة الأربعة أخماس على وجود تأثير سلبي للاختبار على جماعة بشرية ما عندما تنخفض نسبة اختيار هذه الجماعة عن ٨٠٪ (أربعة أخماس) من المجموعة الأكثر اختياراً لهذه الوظيفة أو النوع من العمل.

ويتعين على مؤسسات الأعمال الصناعية المختلفة إثبات صدق الإجراءات أو الاختبارات التي تستخدمها عند اختيار موظفيها. وهناك أدلّة عديدة مقبولة طبقاً للجنة فرص العمل المتكافئة (EEOC) ويتضمن ذلك نتائج دراسات صدق المحتوى أو الصدق المرتبط بالمحك أو صدق المفهوم للاختبار. إلا أن النسخة الأخيرة من توجيهات اللجنة لا توضح النسبة المطلوبة لمعاملات صدق الاختبار، كما أنها لا توضح المعنى للمحكات المرتبطة بالعمل والتي يمكن المصادقة على الاختبارات على أساسها.

تعميم الصدق

لقد تم التأكيد بشكل تقليدي على أن الاختبارات المصادق عليها في حالة معينة دائماً لا تحتفظ بصدقيتها عند استخدامها في حالات أخرى. وعليه فقد درج العاملون في مجال القياس على اعتبار صدق الاختبارات مرتبطاً بالموقع والموقف اللذان يتم فيهما إجراء الاختبار وقد ترتب على هذا، ولسنوات طويلة، إعادة التحقق من صدق الاختبار عند استخدامه في مؤسسة جديدة أو موقع مختلف. إلا أن الأبحاث التي قام بها شميدت وزملاؤه أثبتت أنه يمكن استخدام

اختبارات اختيار الموظفين في عدد كبير من المواقف والمواقع دون المساس بصدق الاختبار (هنتر وشميدت، ١٩٩٠؛ شميدت وهنتر، ١٩٧٧؛ شميدت وهنتر وبيرلمان، ١٩٨٢؛ شميدت وهنتر وبيرلمان وغيرهم، ١٩٨٥). وترتفع فرص بقاء الاختبار صادقاً على الرغم من اختلاف عينة الموظفين والظروف العامة للعمل عندما يتم استخدام اختبارات تعتمد على عينات من المهام الوظيفية التي سيطلب من الممتحن أداؤها في حال حصوله على الوظيفة. كما تشير بعض الدلائل إلى أن جميع اختبارات القدرات الإدراكية هي أدوات تنبؤ صادقة بالأداء في مختلف مجالات العمل.

نظرية الإدراك والمعلومات

لطالما كان علم النفس التطبيقي وعلم القياس النفسي (سايكومتركس) علمين منفصلين انفصالاً تاماً في غرفة الصف الدراسي والمختبر والمكتب، إلا أنه وخلال العقود القليلة الماضية هيأ البحث في مجال علم النفس الإدراكي الفرصة للتقارب بين هذين الفرعين من علم النفس، حيث كثرت في الآونة الأخيرة الأبحاث في العلاقة بين الاستعداد الفردي والعلاج النفسي على الرغم من أن النتائج العملية لهذه الأبحاث حتى اليوم غير مبشرة (سنو، ١٩٩٢). وقد مهدت الدراسات المخبرية لمكونات وارتباطات عملية التعلم وعملية حل المشكلات وغيرهما من العمليات الإدراكية الطريق إلى إعداد اختبارات استعداد لها أساس نظري أكثر ثباتاً. كما ساهمت الدراسات التي تناولت تحليل العمليات الإدراكية التي تتدخل في القدرات اللفظية والمكانية وقدرات الإستدلال والإستقراء في تدعيم أسس بناء الاختبارات في هذه المجالات. ويبقى تحديد ما إذا كانت هذه الاختبارات أكثر قدرة على التنبؤ بالأداء وعلى فهم العمليات الإدراكية التي ينطوي عليها هذا الأداء أمراً الزمن كفيل بإثباته أو نفيه. وفي الوقت نفسه يظهر أنه من الحكمة أن لا يتم استبعاد مقاييس الاستعدادات الخاصة التي يحكم عليها مع الزمن وإلقائها في كومات القمامة (بليغرينو، ١٩٩٢، ١٩٩٤).

الخلاصة

تُعنى الاختبارات التحصيلية بتقييم المعارف والمهارات التي يكتسبها الفرد عن طريق التعليم والتدريب، وتُعنى اختبارات الاستعداد بالتنبؤ بأداء الأفراد في المجالات الأكاديمية والوظيفية. وتتألف الاختبارات التحصيلية من فقرات اختبارية تركز على المادة العلمية المحددة والرسمية التي تتكون منها البرامج التدريبية والمقررات المدرسية بينما تتألف اختبارات الاستعداد من فقرات اختبارية تركز على التجارب العامة وغير الرسمية والمهارات التي تُكتسب عن طريق الصدفة. ومن المفيد لنا هنا أن ننظر إلى اختبارات التحصيل والاستعداد على أنها أجزاء من حقلٍ تعليمي تحتل الاختبارات التي تقيس القدرات الشديدة الخصوصية والمرتبطة بمقرر أو وظيفة ما أحد أطرافه، بينما تحتل الطرف المقابل منه الاختبارات العامة غير اللغوية والعادلة حضارياً. وتختلف اختبارات الاستعداد والاختبارات التحصيلية من حيث الوسط الذي نشأت فيه، فعلى حين تعود أصول الاختبارات التحصيلية إلى المدارس والمؤسسات التعليمية، تُعزى نشأة اختبارات الاستعداد إلى الأوساط الوظيفية.

تم إعداد اختبارات استعداد تقيس استعداداً مهنياً معيناً أو تستخدم مع اختبارات استعداد أخرى كجزء من بطارية اختبار تقيس عدداً من الاستعدادات. وتعد اختبارات الاستعداد النفسية – الحركية والمكانية والميكانيكية والمكتبية من اختبارات الاستعداد الواحد الأكثر شيوعاً. كما تتوفر اختبارات استعداد واحد في مجالات مهارات الحاسب والفن والموسيقا واللغات الأجنبية.

ويعد اختبار سترومبيرغ للبراعة أحد الأمثلة على الاختبارات التي تقيس المهارة الحركية العامة، بينما يعد اختبار كروفرد للبراعة في التعامل مع الأجزاء الصغيرة مثالاً على الأدوات التي تقيس المهارات الحركية التفصيلية والبارعة. ويقيس اختبار البراعة في استخدام الأدوات اليدوية واختبار مينيسوتا للبراعة

اليدوية القدرات على القيام بحركات يدوية مجملة أو دقيقة. ويؤثرالتدريب والممارسة على درجات اختبارات القدرات النفسية - الحركية بشكل كبير وتعد هذه الاختبارات أقل ثباتاً من الاختبارات التي تقيس القدرات الإدراكية.

وتتمتع اختبارات القدرات المكانية كاختبار مينيسوتا للعلاقات المكانية واختبار مينيسوتا التحريري المعدل للوح الأشكال بنسبة صدق جيدة من حيث قدرتها على التنبؤ بأداء الأفراد في المقررات والمهن التي تتطلب مهارات ميكانيكية ومكانية. وكما نلاحظ من اختبار بينيه للإستيعاب الميكانيكي واختبار SRA للاستعداد الميكانيكي، تضم القدرة الميكانيكية عناصر إدراكية بالإضافة إلى عناصر حسية -إدراكية ونفسية- حركية.

ومن أبسط وأقصر الاختبارات التي تقيس القدرات الكتابية اختبار مينيسوتا للمهارات الكتابية (MCT) والذي يتألف من قوائم تضم أسماء وأرقام يطلب من المتحنَين التأكد من خلوها من الأخطاء. وتعد بطارية اختبار القدرات الكتابية واختبار المهام الكتابية العام أدوات قياس للمهارات الكتابية الأكثر تعقيداً. كما تقيس بطارية اختبار استعداد مستخدمي الحاسب وبطارية اختبار استعداد مبرمجي الحاسب واختبار المهارات CRT قدرات الأفراد على برمجة واستخدام الحاسبات الآلية والتي ترتبط إلى حدٍ كبير بالقدرات الكتابية.

تحتل اختبارات بطارية القوات المسلحة للاستعداد المهني (ASVAB) المرتبة الأولى بين اختبارات الاستعدادات المتعددة من حيث عدد الأفراد الذي يتم اختبارهم عن طريقها كل عام، كما تستخدم اختبارات الاستعداد التبايني (DAT) وبطارية اختبار الاستعداد العام (GATB) بشكل واسع. وعلى حين تستخدم بطارية DAT لأغراض التوجيه الأكاديمي، يتم استخدام اختبارات GATB في مجالات التوجيه المهني والحرفي. وغالباً ما تتطلب اختبارات الاستعداد وقتاً طويلاً من الممتحَنين إلا أن عملية مسح استعدادات الموظفين

وبطارية الاستعداد المتعدد الأبعاد يتطلبان وقتاً أقصر من اختبارات ASVAB وDAT وGATB.

وتسمح عملية قياس الاختبارات التي تتألف منها بطارية استعداد ما على نفس العينة من الناس بإجراء مقارنات لدرجات الفرد على أقسام البطارية المختلفة ولدرجات هذا الفرد مع غيره من الممتحَنين. ويجب توخي الحذر عن تفسيرالفروق بين نتائج الفرد في أقسام البطارية المختلفة واعتبارها دالة بشكل إحصائي فقط في حال كون هذه الفروقات أكبر من الخطأ المعياري لقياس فروق الدرجات. ويؤدي قصر اختبارات الاستعدادات المتعددة إلى إنخفاض نسبة ثباتها مما يؤدي بدوره إلى ارتفاع الخطأ المعياري للقياس بالمقارنة مع اختبارات الاستعداد الواحد.

كما تنخفض درجة الصدق التنبؤي لاختبارات الاستعدادات المتعددة، إلا أنه يمكن لدرجات هذه الاختبارات أن تساهم في التنبؤ بمحكات الأداء عند استخدامها مع عدد من مقاييس القدرة الأخرى بالإضافة إلى درجات الفرد على مقاييس الاهتمام والدافع والأداء السابق. ويتم الحكم على اختبارات الاستعداد من حيث فائدتها ليس بالنظر إلى صدقها فحسب بل بالتأكد من أن هذه الاختبارات عادلة وغير منحازة لفئة عرقية معينة. ويتم إثبات عدالة الاختبارات التي تستخدم في اختيار الموظفين والطلاب عن طريق إثبات عدم وجود تأثير سلبي للاختبار على جماعة عرقية أو إثنية معينة، أي عندما لا تنخفض نسبة اختيار هذه الجماعة عن ٨٠٪ (أربعة أخماس) من المجموعة الأكثر اختياراً.

وعلى الرغم من أنه ينصح بإعادة التأكد من صدق الاختبار عند استخدامه لأغراض الاختيار والتصنيف في موقع أكاديمي أو مهني جديد، غالباً ما تتمتع اختبارات الاستعداد بنسبة صدق تعميمي عالية عند استخدامها في مواقع ومواقف متشابهة.

مصطلحات للمراجعة

راجع معاني المصطلحات المدرجة أدناه والتي تم استخدامها في هذا الفصل. الرجاء مراجعة فهرس التعريفات أو المعجم في حال وجود أي التباس حول معاني المصطلحات.

القدرة	الرسم البياني الصوتي
التحصيل	أداة قياس الصوت
التأثير السلبي	مدى النطاق
الاستعداد	القدرة الكتابية
بطارية اختبار الاستعداد	القدرات المتعلقة بالحاسب
الأمانة	اختبار نفسي – حركي
قاعدة الأربعة أخماس	المعايرة المبنية على العرق
الصدق المتزايد	ثبات فروق الدرجات
القدرة الميكانيكية	القدرة المكانية
المتغير المعتدل	الخطأ المعياري لقياس تباين الدرجات
بطارية الاستعدادات المتعددة	
طريقة الحدود الفاصلة المتعددة	تحيز الاختبار
طريقة الانحدار المتعدد	عدالة الاختبار
نمط القدرة المهنية (OAP)	صدق التعميم
بطاقات الألوان المتشابهة	عائلات المهن البصرية

المراجع المقترحة

Gatewood, R., & Perloff, R. (1990). Testing and industrial application. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (2nd ed., pp. 486–501). New York: Pergamon Press.

Greenlaw, P. S., & Jensen, S. S. (1996). Race norming and the Civil Rights Act of 1991. *Public Personnel Management, 25*(1), 13–24.

Guion, R. M. (1991). Personnel assessment, selection, and placement. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed.) (Vol. 2, pp. 327–397). Palo Alto, CA: Consulting Psychologists Press.

Hackett, G., & Watkins, C. E. (1995). Research in career assessment: Abilities, interests, decision making, and career development. In W. B. Walsh & S. H. Osipow (Eds.), *Handbook of vocational psychology: Theory, research, and practice* (2nd ed., pp. 181–215). Mahwah, NJ: Erlbaum.

Hansen, C. P., & Conrad, K. A. (Eds.). (1991). *A handbook of psychological assessment in business*. New York: Quorum Books.

Lubinski, D., & Dawis, R. V. (1992). Aptitudes, skills and proficiencies. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed.) (Vol. 3, pp. 1–5). Palo Alto, CA: Consulting Psychologists Press.

Reschley, D. J. (1990). Aptitude tests in educational classification and placement. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (2nd ed., pp. 148–172). New York: Pergamon Press.

Schmidt, F. L., Ones, D. S., & Hunter, J. E. (1992). Personnel selection. *Annual Review of Psychology, 43*, 627–670.

Snow, R. E. (1992). Aptitude theory: Yesterday, today, and tomorrow. *Educational Psychologist, 27*(1), 5–32.

٧

اختبارات القدرة العقلية العامة

لقد لوحظ ومنذ القدم وجود اختلافات في القدرات الادراكية بين الأفراد والجماعات، ولكن لم تتركز الجهود لقياس هذه الاختلافات ولتحديد أصولها ونتائجها إلا خلال القرن العشرين. منذ عصر أفلاطون وما بعده وضع الفلاسفة نظريات وألفوا في مجالات التعليم و التفكير وطبيعة "العقل" كمقدرة بشرية، ولقد تم تكريس مجال منفرد من علم الفلسفة لدراسة عملية " اكتساب المعرفة" و الشروط التي تنظمها. وعلى الرغم من ذلك لم يتمكن علماء البيولوجيا من غزو هذا المجال وبعض المجالات الأخرى التي كانت سابقاً موضع الاهتمام الرئيسي للفلاسفة حتى القرن التاسع عشر. ولقد أعطت الأبحاث التي قام بها جوهانس مولر وهـ. ل. ف. فون هيلمهولتس وبيير فلورنس وغيرهم من علماء وظائف الأعصاب، الخلفية اللازمة لفهم الإحساس والإدراك الحسي وعمل الدماغ مما قاد إلى نشأة بيئة علمية موائمة لتطوير هذا الفرع من علم النفس.

أدى التدريب الفلسفي والبيولوجي الذي تلقاه علماء النفس الأوائل إلى اتجاه هؤلاء العلماء إلى دراسة الإحساس والإدراك الحسي وطبيعة الوعي وأقسام الدماغ وإلى اعتمادهم على طرق بحث مستوحاة من الفلسفة وعلم

الأحياء كالتأمل الباطني والملاحظة الموضوعية وغيرها. ويعتبر فرانسيس غالتون أحد علماء النفس الرواد في مجال البحث النفسي، إذ قام غالتون، متابعاً لجهود قريبه تشارلز داروين في البحث في أصول الفروقات بين أصناف الأحياء، بدراسة الفروقات الفردية بين المنتمين إلى صنف واحد من الأحياء. ويعتبر غالتون أول من ابتدع أسلوب الاستبانات وسلالم التقدير وتداعي الكلمات وغيرها من الإجراءات المستخدمة اليوم في مجال البحث النفسي. وقد كان البحث في مجال الفروقات في القدرات العقلية بين الأفراد، أو ما أطلق عليه غالتون اسم الذكاء، هو الهدف الرئيسي لابتداع هذه الوسائل.

وعلى الرغم من أن كلمة الذكاء لم تكن مستخدمة بشكل واسع في عصر غالتون، إلا أن تداول هذه الكلمة ازداد بشكل ملحوظ في القرن اللاحق. ولا يعتبر غالتون "أب اختبارات الذكاء" (إذ يحمل هذا اللقب ألفرد بينيه) بالرغم من أبحاثه العديدة والدقيقة في مجال قياس الذكاء والتعرف على الأسس الجينية للعبقرية. وأما الافتراض الذي تقدم به غالتون هو:

إن المعلومات المتعلقة بالأحداث الخارجية لا يمكن أن تصل إلينا إلا عبر الحواس، وأنه بازدياد قدرة حواسنا على الإدراك الحسي للتغيرات المحيطة تزداد خيارات التصرف التي يتوصل إليها ذكاؤنا وحكمنا (غالتون، ١٨٨٣، ص٢٧).

لم يكن افتراضاً خاطئاً كما أثبتت أبحاث حديثة في هذا المجال (جينسين، a١٩٨٢، b١٩٨٢). ولم تكن الاختبارات الحسية - الحركية التي أعدها غالتون (التمييز السمعي والبصري، زمن رد الفعل، تمييز الأوزان، القوة العضلية، الإحساس بالألم واللمس، تذكر الأشكال البصرية، القدرة على التنفس، حجم الرأس وغيرها) ذات فائدة تذكر في التنبؤ بأداء الأفراد للمهام المعقدة والتي يفترض أن الذكاء الفردي يلعب دوراً في إنجازها (ويسلر، ١٩٠١). وأما اختبارات ألفرد بينيه وثيوفيل سايمون ذات التوجه المدرسي والتي شكلت جزءاً من

اختبارات الذكاء (١٩٠٥، ١٩٠٨، ١٩١١) التي أعداها، فقد أثبتت قدرتها على التنبؤ بالتحصيل الأكاديمي/المدرسي مما أدى اعتماد مقياس بينيه – سايمون للذكاء ومقياس ستانفورد –بينيه للذكاء (وهو النسخة الأمريكية المترجمة والمعدلة عن مقياس بينيه– سايمون) كمحك أساسي تقارن جميع اختبارات الذكاء به.

يعتبر مقياس بينيه – سايمون مقياساً عمرياً كما نلاحظ من فقرات الاختبار في الجدول ٧–١ والتي تم أخذها من نسخة عام ١٩١١ من المقياس، إذ يتم أول الأمر عند إجراء هذا الاختبار تحديد العمر القاعدي basalage للممتحَن وهو أعلى مستوى عمري تمكن المتحَن من اجتياز جميع الاختبارات الفرعية المرتبطة به بنجاح، ثم يعطى الممتحَن الاختبارات التي تنتمي إلى المراحل العمرية اللاحقة الواحد بعد الآخر حتى يتم الوصول إلى السقف العمري Ceiling age وهو العمر الذي فشل الممتحن في النجاح بأي من الاختبارات الفرعية المرتبطة به. وكانت عملية وضع درجات هذا الاختبار تتم عن طريق منح الطالب درجة كاملة إذا اجتاز جميع الاختبارات الفرعية لعمر ما بنجاح وأجزاء من الدرجة في حال اجتيازه لبعض الاختبارات الفرعية لعمر معين. يتم بعد ذلك جمع الدرجات للتوصل إلى العمر العقلي للممتحَن mental age. ولما كان الهدف الرئيسي لاختبارات بينيه – سايمون هو التعرف وبشكل موضوعي على الطلاب الذين لن يحصلوا على أية فائدة علمية من البقاء في الفصول الدراسية مع أقرانهم، كان يتم تصنيف الأطفال الذين يتأخر عمرهم العقلي عن عمرهم الفعلي بشكل كبير كمتخلفين عقلياً ثم يتم إلحاقهم ببرامج التربية أو التعليم الخاص.

الجدول ٧-١

اختبارات من نسخة عام ١٩١١ لمقياس سايمون - بينيه

عمر ثلاث سنوات:

- يشير إلى أنفه وعينيه وفمه
- يردد عددين
- يرقم الأشكال في صورة
- يعرف كنيته
- يردد جملة من ستة مقاطع

عمر أربع سنوات:

- يعرف نوعه (ذكر أو أنثى)
- يتعرف على السكينة والمفتاح وقطعة النقود
- يردد ثلاثة أعداد
- يقارن خطين

عمر خمس سنوات:

- يقارن وزنين
- ينقل رسم المربع
- يردد جملة من عشرة مقاطع
- يعد أربعة قطع نقدية
- يضم قسمي مثلث مقسوم

عمر ست سنوات:

- يميز بين الصباح وبعد الظهيرة
- يعرف المصطلحات الشائعة من حيث استخدامها
- ينقل رسم المضلع

- يعد ١٢ قطعة نقود
- يميز بين الصور التي تحوي وجوه قبيحة وجميلة

عمر سبع سنوات:

- يدل على يده اليمنى وأذنه اليسرى
- يصف صورة
- ينفذ ثلاث أوامر تعطى إليه في نفس الوقت
- يعد قيمة ستة قطع نقدية ثلاثة منها مضاعفة
- يسمي أربعة من الألوان الأساسية

عمر ثماني سنوات:

- يقارن شيئين موجودين في الذاكرة
- يعد من عشرين إلى صفر
- يلاحظ الأشياء المفقودة من الصور
- يعرف اليوم والتاريخ
- يردد خمسة أعداد

عمر تسع سنوات:

- يرجع الباقي من قطعة نقدية قيمتها ٢٠
- يعرف الأشياء الشائعة بالاعتماد على ما هو أعمق من استخدامها
- يتعرف على كافة أنواع القطع النقدية
- يسمي الأشهر بالترتيب
- يجيب على الأسئلة البسيطة ويستوعبها

عمر ١٠ سنوات:

- يرتب خمسة مكعبات تبعاً لأوزانها

- يعيد رسم أشياء من الذاكرة
- ينتقد العبارات السخيفة
- يجيب على الأسئلة الصعبة أو يستوعبها
- يستخدم ثلاث كلمات ضمن ما لا يزيد عن جملتين

عمر ١٢ سنة:

- يعارض ما يقال عن أطوال الخطوط
- يركب جملة تضم ثلاث كلمات معطاة
- يسمي ستين كلمة خلال ٣ دقائق
- يعرف ثلاث كلمات مجردة
- يفهم معنى جملة كلماتها مبعثرة

عمر ١٥ سنة:

- يردد ٧ أرقام متتابعة
- يجد ثلاث كلمات على نفس وزن كلمة معطاة خلال دقيقة واحدة
- يردد جمل من ٢٦ مقطعاً
- يفسر الصور ويحللها
- يفسر الحقائق المعطاة

البالغ:

- يحل اختبار قص الأوراق
- يعيد ترتيب أجزاء مثلث في الخيال
- يوضح الفرق بين أزواج من المفاهيم المجردة
- يعطي ثلاث فروق بين الملك ورئيس الجمهورية
- يحدد الأفكار الرئيسة في نص مختار

تعاريف وتصاميم وتطبيقات

الذكاء المعرّف وغير المعرّف

من الغريب جداً أنه وبعد قرن كامل من الأبحاث في مجال الذكاء لا يزال علماء النفس يختلفون حول التعريف الدقيق لمفهوم الذكاء، فتعريف بينيه للذكاء على أنه "القدرة على إطلاق الأحكام الجيدة واستيعاب الأمور بشكل جيد والتفكير بشكل جيد" (بينيه وسيمون، ١٩٠٥، ص١٩٦) يناقض التعريف الشائع للذكاء على أنه، وببساطة، القدرة على التعلم. كما تتوفر تعريفات أخرى للذكاء كتعريف لويس تيرمان له على أنه القدرة على "التفكير التجريدي" abstract thinking، والتعريفات البيولوجية التي ترى الذكاء على أنه "قدرة التكيف مع البيئة" أو "إمكانية تعديل الجهاز العصبي" (ردولف بينتنر) ("الذكاء" ١٩٢١). وعلى الرغم من كثرة التعاريف تبقى جميعها مرتبطة بتأثيرات، لا بأصول، الذكاء. ومن الجائز أن تقود الأبحاث المعاصرة في مجال العمليات الإدراكية إلى تعريف أكثر دقة للذكاء، وفي انتظار أن يتحقق هذا يمكن الاعتماد على تعريف إي. جي. بورينغ للذكاء على أنه "كل ما تقيسه اختبارات الذكاء" والذي يعد تعريفاً عملياً ومنطقياً. وعلى الرغم من تشابه اختبارات الذكاء إلا أن الواحد من هذه الاختبارات لا يمكن أن يستخدم عوضاً عن الآخر وذلك بسبب الاختلافات الكبيرة في محتويات الاختبارات وطرق إعدادها وبنائها وإجرائها. لهذا السبب يجب علينا عند إعطاء الأفراد نتائجهم في اختبارات الذكاء أن نحدد الاختبار الذي بنيت عليه هذه النتائج.

تصميم الاختبارات

بالرغم من أن الهدف العام لجميع اختبارات الذكاء هو قياس القدرة العقلية للأفراد، تختلف هذه الاختبارات بشكلٍ ملحوظ من حيث التصميم والإجراء

والتصحيح وتفسير النتائج. ويمكن لفقرات اختبارات الذكاء أن تستند، كما هو الحال مع اختبار ستانفورد - بينيه وسلسلة اختبارات ويكسلر، إلى الخبرات الملموسة والتحليل العملي لأداء الأفراد في مهام إدراكية معينة. ويتم تطبيق عدد من إجراءات تحليل فقرات الاختبار البسيطة أو المعقدة والمتعددة المتغيرات (كالتحليل العاملي) وذلك للحصول على أداة قياس سايكومترية (نفسية) تتألف من مجموعة من الفقرات التي يمكن أن تساهم في قياس القدرات الإدراكية التي يشار إليها بكلمة "الذكاء". كما لعبت نظريات النمو وعلم النفس العصبي دوراً في تصميم بعض اختبارات الذكاء وتعدنا الأبحاث الحالية في مجالات الإدراك وعلم النفس العصبي بمساهمة مشابهة في المستقبل.

ويمكن لاختبار ذكاءٍ ما أن يكون ذو أسس عملية أو نظرية وأن يتألف من فقرات لفظية وغير لفظية أو فقرات تتطلب من الممتحن أن يتعامل مع أشياء ملموسة أو أن يجيب بشكل شفهي أو كتابي على أسئلة محددة. وتحتوي البرامج H-1 إلى H-6 الموجودة على القرص المرفق بالكتاب على نماذج لفقرات اختبارات الذكاء (كلمات مبعثرة، تذكر الأرقام، متسلسلات عددية، مسائل، إيجاد العلاقات والارتباطات، تداعي الأفكار الإبداعي...إلخ).

قد يتألف اختبار الذكاء من أداة قياس ذات مستوى واحد أو من عدد من الأدوات المعدة لأعمار مختلفة (الاختبار المتعدد المستويات)، وقد يخدم الاختبار ذو المستوى الواحد عدداً كبيراً من الأشخاص كالأطفال والبالغين أو الأطفال والبالغين معاً. ومن الممكن ترتيب فقرات اختبارات الذكاء ضمن اختبارات فرعية مختلفة تتشابه فقراتها من حيث الشكل أو المضمون، كما يمكن خلط الفقرات ضمن كتيب اختبار لولبي أو عادي. ويمكن إجراء اختبارات الذكاء لفرد واحد أو لمجموعة كبيرة من الأفراد في ذات الوقت. ويمكن العودة إلى القائمتين ٠ و ١ من برنامج H-7 المرفق بالكتاب للحصول على أسماء لاختبارات ذكاء فردية وجماعية وعلى اقتراحات لتصميم وإجراء وتصحيح هذه الاختبارات.

على الرغم من أن تصحيح الاختبارات الفرعية ضمن اختبار ذكاء ما يتم بشكل منفصل عن الاختبارات الفرعية الأخرى، من الممكن جمع درجات الاختبارات الفرعية لتشكل جزءاً من المجموع النهائي أو المجموع النهائي لدرجات الاختبار. ويعتمد تصحيح الاختبارات على درجة موضوعية الأسئلة المستخدمة، ففي حال كون هذه الأسئلة موضوعية تماماً ولا تتطلب من المصحح أن يحكم على الإجابة، تكون درجات الاختبار موضوعية وخالية من الذاتية. أما في حال الطلب من المصحح إصدار حكم على إجابة الممتحَن من حيث الدقة والأسلوب فلا تعد درجات الاختبار موضوعية تماماً.

يتم إجراء معظم اختبارات الذكاء باستخدام الطريقة التقليدية التي تطلب من الممتحنين الإجابة على كافة أسئلة الاختبار، إلا أنه (وحتى في الاختبارات العريقة كاختبار ستانفورد – بينيه وسلسلة اختبارات ويكسلر)، يمكن تجاوز عدد من الأسئلة إذا كان الاعتقاد المرجح هو أن الممتحن يعرف الإجابة إذا ما طرح السؤال عليه. وقد كثر في السنين الأخيرة استخدام الاختبارات التكيفية المحوسبة (CAT) والتي تكيف الاختبارات الفرعية التي تُعطى للطالب تبعاً لإجاباته (راجع الفصل الثالث). ومن الممكن أن يقوم الحاسب كذلك بتفسير النتائج وتحليلها وإعداد التقرير المناسب للدرجات.

يُنصح مصممو الاختبارات بإتخاذ القرارات المتعلقة بإجراء وتصحيح وتفسير نتائج الاختبارات قبل بناء الاختبار، وفي حال كشفت التجارب العملية للاختبار عن الحاجة إلى إجراء تعديلات في طريقة وضع الدرجات أو إجراء الاختبار يتعين على الجهة المسؤولة عن الاختبار البحث في ما إذا كانت التعديلات بسيطة تجمّل أو تحسّن الاختبار أو ما إذا كان هنالك حاجة إلى إعداد شكل جديد من الاختبار.

التطبيقات

تُستخدم اختبارات الذكاء لتصنيف الطلاب المتخلفين أو الموهوبين عقلياً، إلا أن لهذه الاختبارات استخدامات أخرى. فغالباً ما تجرى اختبارات الذكاء للأطفال ذوي المستويات العلمية المتدنية في المدارس وذلك لمعرفة ما إذا كان ضعف أو تأخر القدرات العقلية للطالب هي السبب وراء تدني تحصيله. كما يمكن مقارنة درجات الطالب في اختبارات الذكاء مع درجاته في اختبارات التحصيل وذلك لتحديد ما إذا كان هنالك فجوة لدى هذا الطالب بين قدرته وتحصيله، ويعني هذا أن إمكانيات الطالب تفوق تحصيله الفعلي مما قد يقودنا إلى التعرف على أسباب تدني تحصيل الطالب والتي يمكن أن تكون أسباب عاطفية أو اجتماعية أو غيرها من الأسباب التي لا ترتبط بقدرات الطالب الإدراكية.

تُستخدم بعض اختبارات الذكاء الجماعية بشكل حصري في مجال اختيار وتسيب الموظفين (كاختبار وندرليك للموظفين)، كما تستخدم اختبارات الذكاء في تحديد الأفراد الذين يتلقون منحاً دراسية أو يرغبون في الإنضمام إلى بعض المنظمات مثل منظمة منسا[١] (Mensa) ولا تقتصر استخدامات اختبارات الذكاء على المواقع الأكاديمية والمهنية، إذ تستخدم المستشفيات النفسية وعيادات إرشاد الأطفال وغيرها العديد من اختبارات الذكاء الفردية لأغراض تشخيصية وعلاجية للحالات النفسية. كما تجدر الإشارة كذلك إلى الاستخدامات الواسعة لاختبارات الذكاء (والتي تشهد عليها ملخصات العديد من أطروحات الدراسات العليا والمقالات العلمية) في الأبحاث التي تتناول الفروقات الفردية والجماعية

(١) وهي منظمة عريقة تضم الأفراد ذوي معدلات الذكاء الأعلى من الطبيعية. (المترجم).

في القدرات الإدراكية وفي غيرها من القضايا النفسية ذات الأهمية في عالمنا المعاصر.

سيوفر ما بقي من هذا الفصل للقارئ وصفاً عاماً لأكثر اختبارات الذكاء الفردية والجماعية (أو اختبارات القدرة الإدراكية العامة) تداولاً. ويوضح الجدول ٧-٢ أسماء اختبارات الذكاء العشرين الأكثر ذكراً في الكتب المنشورة بين ١٩٩١ و١٩٩٧، أما النمو الإدراكي وإضطراباته فسيتم بحثها في الفصل الثامن مع المزيد من الحديث عن اختبارات الذكاء.

اختبارات الذكاء الفردية

على الرغم من تعدد الطرق التي يمكن أن نصنف فيها اختبارات الذكاء، يعد التصنيف الذي يستند إلى طريقة إجراء الاختبارات -جماعية أو فردية- التصنيف الأكثر بساطة. وقد أجريت اختبارات الذكاء الأولى من قبل شخص واحد ولممتحَن واحد، ويعتبر الكثير من علماء النفس هذا الإجراء الطريقة الوحيدة الناجحة لقياس القدرة الإدراكية العامة للأفراد، إذ تساعد صفة الود والتواصل المباشر التي تتميز بها هذه الطريقة على تحفيز الممتحَنين للإجابة على الأسئلة بالشكل الأفضل، كما تتيح الاختبارات الفردية الفرصة للمراقب أن يدرس سلوك الممتحَن أثناء الاختبار مما يساعد في عمليات التشخيص النفسي وفي تحديد درجة صدق الاختبار.

الجدول ٧-٢

اختبارات الذكاء العشرين الأكثر ذكراً في كتب علم النفس المنشورة بين ١٩٩١ و١٩٩٧

اسم اختبار الذكاء	عدد المرات التي ذكرت فيها كتب علم النفس الاختبار
مقاييس بايلي لنمو الرضع	١٣٠
اختبار القدرات الإدراكية	٢٥
اختبار الذكاء العادل حضاريا	٢٤
مقاييس القدرات التباينية	١٨
اختبار كوفمان لذكاء المراهقين والبالغين	١٩
بطارية كوفمان لتقييم الأطفال	٨٥
اختبار كوفمان المختصر للذكاء	٢٢
اختبار مصفوفة المتشابهات	١٢
مقياس ماكارثي لقدرات الأطفال	٥٥
اختبار أوتيس – لينون للقدرة المدرسية	١٦
اختبار "بيبودي" المعدل للمفردات المصورة	١٦٨
مصفوفات رافن المتدرجة	١٥٢
مقياس معهد شيبلي لأسلوب المعيشة	٢٦
اختبار سلوسون للذكاء المعدل	١٩
اختبار ستانفورد – بينيه للذكاء (جميع الطبعات)	١٣٠
اختبار الذكاء اللالفظي	٢٤
مقياس ويكسلر لذكاء البالغين (جميع الطبعات)	٩٣١
مقياس ويكسلر لذكاء الأطفال (جميع الطبعات)	٦٨٩
مقياس ويكسلر لذكاء الأطفال دون عمر المدرسة وفي المرحلة الإبتدائية (جميع الطبعات)	٩٨
اختبار وندرليك للموظفين	٩

ويعتبر الكثيرون نتائج اختبارات الذكاء الفردية أكثر ثباتاً من غيرها من الاختبارات وذلك بسبب السيطرة الكاملة التي يتمتع بها المشرف على ظروف الاختبار والتي تحفز الطالب على الإجابة بشكل صادق وواعي على أسئلة الاختبار. إلا أنه من المهم أن نذكر أن اختبارات الذكاء الجماعية لا تحتوي على الأسئلة المفتوحة الإجابة التي تضمها اختبارات الذكاء الفردية مما يعني أن عملية وضع الدرجات لاختبارات الذكاء الجماعية تتميز بالموضوعية مما يجعلها على نفس القدر من الثبات الذي تتميز به الاختبارات الفردية.

مقياس ستانفورد - بينيه للذكاء

كما يمكن أن نلاحظ من اسم الاختبار، كان اختبار ستانفورد - بينيه الأساسي اختباراً تم تعديله وإعادة تقنينه في الولايات المتحدة عن النسخة الأصلية من مقياس بينيه - سيمون للذكاء. ولطالما كانت الطبعات الثلاثة الأولى من اختبار ستانفورد - بينيه (١٩١٦، ١٩٣٧، ١٩٦٠) المقاييس الأكثر استخداماً في الولايات المتحدة لتقييم القدرات الذهنية للأفراد، إلا أن السنوات الأخيرة شهدت تراجعاً في استخدام مقياس ستانفورد - بينيه لصالح مقاييس ويكسلر للذكاء (انظر الجدول ٧-٢). وقد اتبعت النسخ الأولى من اختبار ستانفورد - بينيه طريقة التطبيق ذاتها التي اتبعها اختبار بينيه - سيمون والتي تعتمد على تحديد العمر القاعدي للممتحَن ثم الاستمرار باختبار الممتحَن في جميع مستويات المراحل العمرية حتى يصل إلى مرحلة يفشل فيها في جميع الاختبارات الفرعية المرتبطة بعمر ما (وهذا ما أشرنا إليه سابقاً باسم سقف الامتحان). يتم بعد ذلك حساب العمر العقلي للممتحن بالأشهر والسنوات عن طريق إعطاء درجات لكل اختبار فرعي نجح فيه أو افترض نجاحه فيه (هذا في حال وقوع هذا الاختبار الفرعي دون العمر القاعدي للممتحن).ويتم آخر الأمر حساب معدل ذكاء الممتحن Intelligence quotient,IQ عن طريق تقسيم عمر الممتحن العقلي (بالأشهر) على عمره الحقيقي (بالأشهر) ثم ضرب الناتج بالعدد ١٠٠، وقد كان المعدل العام للأفراد في الولايات المتحدة ١٠٠ مع نسب ضئيلة تحت أو فوق هذا المعدل.

تم استخدام النسخ الثلاثة الأولى من اختبار ستانفورد - بينيه بشكل كبير في المدارس وذلك للتعرف على الطلاب المتخلفين عقلياً والذين يحتاجون إلى صفوف تربية خاصة. كما استُخدم الاختبار في العديد من الأبحاث والدراسات والأغراض العملية، وتم اعتباره المقياس الذي تقارن به جميع التطورات في مجال دراسة وقياس الذكاء.

وتعد النسخة الأحدث من اختبار ستانفورد - بينيه (الطبعة الرابعة SB-IV) نسخة محسّنة (إلا أنها لا تخلو من العيوب) عن الاختبار، استندت إلى أسس نظرية أكثر ثباتاً وتم بناؤها وفق طرق أكثر دقة من سابقاتها. وكما يوضح الشكل ٧-١، تنظر الأسس النظرية لاختبار SB-IV إلى نموذج الذكاء كنموذج هرمي يحتل الذكاء العام g المستوى الأعلى فيه، ثم تأتي ثلاث قدرات عامة في المستوى الثاني وهي: القدرات المتبلورة والقدرات التحليلية الفطرية (المائعة) والذاكرة قصيرة الأمد، أما المستوى الثالث فتحتله متغيرات الاستدلال الثلاثة: اللفظية والكمية والتجريدية/البصرية.

يتم تقييم متغيرات الاستدلال الثلاثة ومتغير الذاكرة القصيرة الأمد في اختبار SB-IV عن طريق ١٥ اختباراً موزعة على النحو التالي:

- يتم تقييم الاستدلال اللفظي عن طريق اختبار المفردات والاستيعاب والتعرف على التفاهات (أو ما لا معنى له) واختبار العلاقات اللفظية؛
- يتم تقييم الاستدلال الكمي عن طريق اختبارات كمية واختبارات المتسلسلات العددية واختبارات بناء المعادلات؛
- يتم قياس الاستدلال التجريدي/البصري عن طريق تحليل النماذج ونسخ الصور والمصفوفات واختبارات ثني وقص الورق؛
- يتم تقييم الذاكرة قصيرة الأمد عن طريق اختبارات الذاكرة التتابعية وتذكر الجمل والأرقام والأشياء.

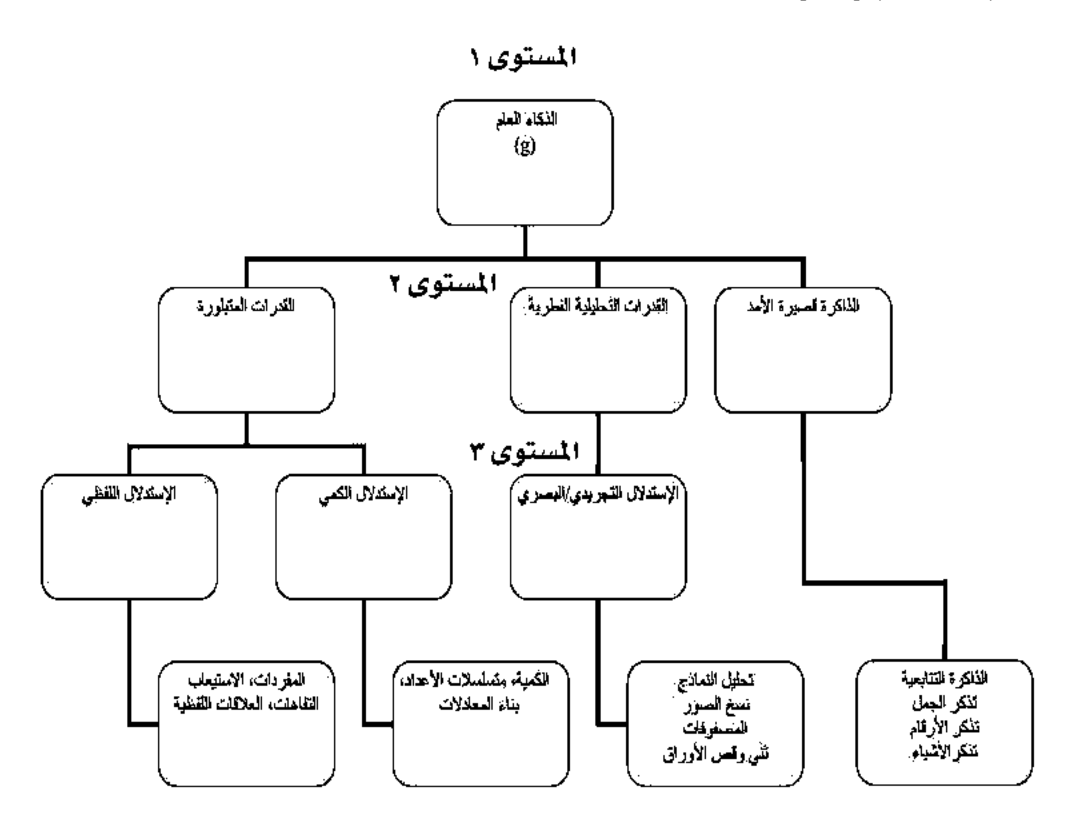
يعد اختبار SB-IV اختباراً تكيفياً يتألف من عدة مراحل ويتم تحديد المرحلة التي يبدأ عندها الفرد عن طريق اختبار للمفردات أو ما يسمى باختبار تحديد الوجهة. ويتم تحديد النقطة التي يبدأ عندها اختبار تحديد الاتجاه بناءاً

على العمر الفعلي للممتحَن وتستمر عملية الاختبار واستخدام أسئلة أكثر صعوبة حتى يفشل الممتحَن في الإجابة على ثلاثة أو أربعة أسئلة في مستويين متلاحقين ويسمى المستوى الأعلى منهما "المستوى الحرج" (أو الفاصل) والذي يحدد مع عمر الممتحَن الحقيقي نقطة البداية للممتحَن في الاختبارات الأربعة عشرة التي تؤلف اختبار SB-IV يتم بعد ذلك إعطاء الممتحن أسئلة تقل صعوبتها تدريجياً حتى يصل إلى الممتحن إلى مستويين أجاب فيهما بشكل صحيح على جميع الفقرات الاختبارية. وتحدد هذه العملية المستوى القاعدي للفرد في كل من الاختبارات الأربعة عشرة. تتلو هذه عملية، عملية إعطاء الممتحَن أسئلة تتزايد صعوبتها تدريجياً حتى يفشل الممتحَن في الإجابة على ثلاثة أو أربعة أسئلة من مستويين متتابعين مما يمكن من تحديد سقف العمر لذلك الاختبار.

وتمثل الإجابات الصحيحة للممتحَن في اختبار معين من اختبارات SB-IV، درجاته الأولية في كل اختبار، ثم يتم تحويل هذه الدرجات إلى درجات عمرية قياسية (SAS) مبنية على الفئة العمرية للممتحن وبوسط حسابي يساوي ٥٠ وانحراف معياري يساوي ٠,٨ يتم حساب الدرجات الخام لكل من الفئات الأربعة المذكورة أعلاه (الإستدلال اللفظي والإستدلال الكمي والإستدلال التجريدي/ النظري والذاكرة قصيرة الأمد) عن طريق جمع درجات الاختبارات الثلاثة أو الأربعة التي تتألف منها كل فئة، ثم يتم بعد ذلك تحويل الدرجات الناتجة إلى درجات مؤشر قياسية (درجات SAS لكل فئة) بوسط حسابي يساوي ١٠٠ وانحراف معياري يساوي ٠,١٦ كما يتم جمع درجات الفئات الأربعة للحصول على معدل درجات قياسي مركب يستند إلى مؤشر وسطه الحسابي ١٠٠ وانحرافه المعياري ١٦ ومداه بين ٣٦ و١٦٤.

تم قياس اختبار SB-IV على عينة من مواطني الولايات المتحدة تراوحت أعمار أفرادهابين عامين وثلاثة وعشرون عاماً وأحد عشر شهراً. وقد تم اختيار أعداد كبيرة من أفراد العينة من فئات ذات مستوى اجتماعي وتعليمي عالٍ مما

أدى إلى مشاكل في وضع درجات هذا الاختبار لم يتم حلها حتى اليوم. ويعاني اختبار SB-IV من مشاكل أخرى متعلقة بعدم تساوي المهارات التي يقيسها الاختبار عند الفئات العمرية المختلفة بالإضافة إلى كون المعلومات المتوفرة عن ثبات الاختبار غير كافية.



الشكل ٧-١

النموذج النظري واختبارات ستانفورد – بينيه النسخة الرابعة (SB-IV)

جميع حقوق الطبع محفوظة لشركة ريفرسايد للنشر، ١٩٨٦، تم النقل بتصرف عن مقياس ستانفورد – بينيه للذكاء، الطبعة الرابعة، دليل إجراء وتصحيح الاختبار، النموذج الهرمي، الصفحة ٤ بعد موافقة شركة ريفر سايد للنشر.

اختبارات WAIS-R و WISC-III و WPPSI-R

تعد سلسلة الاختبارات التي قدمها ديفيد ويكسلر للعالم عام ١٩٣٩ أكثر أدوات قياس الذكاء الفردية شعبية في يومنا هذا، وتتألف السلسلة من ثلاثة اختبارات هي: ١) اختبار ويكسلر المعدل لذكاء البالغين (WAIS-R) للأفراد الذين تتراوح أعمارهم بين ١٦ و ٧٤ عاماً. ٢) الطبعة الثالثة من مقياس ويكسلر لذكاء الأطفال بين أعمار ٦ إلى ١٦,١١ شهراً (WISC-III). ٣) والنسخة المعدلة عن مقياس ويكسلرلذكاء الأطفال دون عمر المدرسة وأطفال المرحلة الإبتدائية (WPPSI-R، لأعمار ٣ إلى ٧,٣ سنة). وتوجد عدة اختلافات بين هذه الاختبارات وبين الطبعات الأربعة من مقياس ستانفورد – بينيه:فعلى حين تم إعداد الاختبار الأول من مقياس ستانفورد – بينيه بهدف استخدامه ضمن سياق المدارس حصراً، تم إعداد أول اختبارات سلسلة ويكسلر (مقياس ويكسلر – بلفيو الشكل الأول) بهدف استخدامه في مجالات العلاج النفسي. وبينما تعد الطبعات الثلاثة الأولى من اختبار ستانفورد – بينيه مقاييس للعمر تم تصنيف الأسئلة فيها ضمن مجموعات تتوافق مع فئات عمرية مختلفة ويمنح المتحَن درجات للإجابات الصحيحة تتمثل بعدد من الأشهر ترفع من العمر العقلي لهذا المتحن، تم تصنيف أسئلة اختبارات ويكسلر ضمن مجموعة من الاختبارات الفرعية التي تحتوي على أسئلة متشابهة في المحتوى ويحصل المتحن في اختبارات ويكسلر على عدد من الدرجات لكل إجابة صحيحة يتوصل إليها. وأخيراً، يتم التعبير عن أداء الفرد في الطبعات الثلاثة الأولى من اختبار ستانفورد – بينيه عن طريق معدل درجات واحد (هو معدل ذكاء نسبي IQ في الطبعتين الأولى والثانية ومعدل ذكاء انحرافي IQ في الطبعة الثالثة). أما الطبعة الرابعة من مقياس ستانفورد – بينيه فتختلف في تصميمها وإجرائها ووضع درجاتها (كما أشرنا في الباب السابق من هذا الفصل) عن الطبعات الأخرى من ذات الاختبار وعن الاختبارات الفرعية في مقياس ويكسلر.

يضم الجدول ٧-٣ عشرة إلى اثني عشر اختباراً فرعياً تؤلف اختبارات WAIS-R وWISC-III وWPPSI-R تم ترتيبها بحسب الترتيب الذي تجرى فيه هذه الاختبارات، وكما تشير الأحرف التي ترمز إلى نوع الاختبار الفرعي (لفظي، أدائي..ألخ) ينتقل المتحَن من اختبار لفظي إلى اختبار أدائي ثم لفظي وهكذا دواليك. وتم إلحاق اختبارين فرعيين إضافيين باختبار WPPSI-R وثلاثة اختبارات فرعية إضافية باختبار WISC-III، ولا تُستخدم هذه الاختبارات الإضافية إلا في حال عدم استخدام أحد الاختبارات الفرعية الأساسية لسبب من الأسباب أو عن حدوث خطأ في إجرائها.

الجدول ٧-٣ مقاييس ويكسلر للذكاء

WPPSI-R (الأعمار ٣٠-٧،٣)	WISC-III (الأعمار ٦-١٦،١١)	WAIS-R (لأعمار ١٦-٧٤)
تجميع الأشياء (أ)	إكمال الصور (أ)	معلومات (ل)
معلومات (ل)	معلومات (ل)	إكمال الصور (أ)
تصميم هندسي (أ)	ترميز (أ)	المدى الرقمي (ل)
الإستيعاب (ل)	التشابهات (ل)	ترتيب الصور (أ)
تصميم المكعبات (أ)	ترتيب الصور (أ)	مفردات (ل)
حساب (ل)	حساب (ل)	تصميم المكعبات (أ)
المتهاهات (أ)	تصميم المكعبات (أ)	حساب (ل)
مفردات (ل)	مفردات (ل)	تجميع الأشياء (أ)
إكمال الصور (أ)	تجميع الأشياء (أ)	الإستيعاب (ل)
التشابهات (ل)	الإستيعاب (ل)	رموز رقمية (أ)
(حيوانات على ألواح مع قبضة مسمارية) (أ)	(البحث عن الرموز) (أ)	التشابهات (ل)
(جمل) (ل)	(المدى الرقمي) (ل)	
	(متاهات) (أ)	

ملاحظة: الاختبارات المكتوبة بين قوسين هي اختبارات إضافية والرمز (أ) يعني اختبار أداء بينما يعني الرمز (ل) اختبار لفظي/لغوي.

يتم ترتيب الاختبارات الفرعية في مقاييس ويكسلر من الأسهل إلى الأصعب وتعتمد نقطة البداية للفرد على عمره الفعلي. تتابع الأسئلة على المتحن في اختبار فرعي ما حتى يفشل في الإجابة على عدد من الأسئلة المتتابعة مما يؤدي إلى انتقال هذا الفرد إلى اختيار فرعي آخر. يجب إتمام جميع الاختبارات

الفرعية ضمن فترات زمنية محددة ويحصل المتحَن على درجات إضافية عند إتمام بعض الأسئلة في وقت أسرع من المطلوب. ويستغرق كل من اختباري WPPSI-R وWAIS-R ٧٥ دقيقة، بينما يستغرق اختبار WISC-III ٥٠ إلى ٧٠ دقيقة.

فيما يلي وصف مختصر للاختبارات الفرعية لمقياس WAIS-R والتي تمثل محتويات كافة مقاييس ويكسلر:

- معلومات: وهي عدد من الأسئلة المرتبطة بالمعرفة العامة والتي يتوجب على المتحَن الإجابة عليها ببضعة كلمات أو باستخدام الأرقام
- إكمال الصور: ويطلب من المتحَن في هذا النوع من الأسئلة التعرف على الأجزاء الناقصة من عدد من الصور الموجودة على بطاقات.
- المدى الرقمي: يتم قراءة عدد من الأرقام بصوت مرتفع أمام المتحَن ثم يطلب منه إعادة ترديد الأرقام من الأصغر إلى الأكبر في المجموعة الأولى ومن الأكبر إلى الأصغر في المجموعة الثانية.
- ترتيب الصور: يُعطى المتحَن عدد من البطاقات التي تضم صوراً ثم يطلب منه إعادة ترتيب الصور لرواية قصة ذات معنى.
- المفردات: يُطلب من المتحَن تعريف عدد من المفردات تتدرج صعوبتها مع تقدم الطالب في الاختبار.
- تصميم المكعبات: يطلب من المتحَن تقليد عدد من التصاميم الهندسية المؤلفة من مربعات حمراء وبيضاء والمرسومة على بطاقات باستخدام مكعبات حمراء وبيضاء.
- الحساب: يُطلب من المتحَن حل عدد من المسائل الحسابية مرتبة بالتدريج من الأسهل إلى الأصعب.

- تجميع الأشياء: يُعطى الممتحن عدد من الصور التي تم قصها وبعثرة أجزائها ويطلب منه إعادة ترتيبها لصنع شيء ذو معنى.
- الإستيعاب: يُطلب من الممتحن الإجابة على عدد من الأسئلة مرتبة بالتدريج من الأسهل إلى الأصعب بشكل مفصل ودقيق.
- الرموز الرقمية :يتم تعبئة صناديق فارغة برموز تناسب الأرقام المعلقة في أعلى الصندوق.
- التشابهات: يُطلب من الممتحَن تحديد وجه الشبه بين شيئين.

بعد إنتهاء الممتحن من كافة الاختبارات الفرعية، يتم تحديد الدرجات الأولية لكل اختبار فرعي ثم يتم تحويل هذه الدرجات إلى مؤشر قياسي للدرجات وسطه الحسابي ١٠ وإنحرافه المعياري ٣. وتضاف درجات المؤشر القياسي لاختبارات الأداء واللغة ثم يتم تحويلها إلى معدلات ذكاء IQ لفظية (لغوية) وأدائية، كما يتم إضافة درجات جميع الاختبارات الفرعية وذلك للحصول على معدل الذكاء IQ العام للممتحَن. ويتم تصنيف معدلات الذكاء الانحرافية الثلاثة، والتي تستند إلى مؤشر للدرجات القياسي وسطه الحسابي ١٠٠ وإنحرافه المعياري ١٥، على النحو التالي:

١٣٠ وما فوق = متفوق للغاية

١٢٠ – ١٢٩ = متفوق

١١٠ – ١١٩ = فوق المعدل

٩٠ – ١٠٩ = على المعدل

٨٠ – ٨٩ = أدنى من المعدل

٧٠ – ٧٩ = على الحدود

٦٩ وما دون = متخلف عقلياً

ويمكن الحصول على معلومات تشخيصية إضافية من الفروق الدالة إحصائياً بين معدلات الذكاء اللغوية والأدائية ودرجات المؤشر القياسي في الاختبارات الفرعية من مقياس ويكسلر بالإضافة إلى عدد من درجات العامل الواحد (الإستيعاب اللغوي، التنظيم الإدراكي – الحسي، وعدم الإكتراث بالموه في اختبار WISC-III) وعلى الرغم من أنه يمكن الحصول على تفسير لنتائج WAIS-Rو WISC-III وWPPSI-R وذلك عن طريق خدمة التصحيح والتفسير المحوسبة التي توفرها الشركة المتحدة لعلم النفس (The Psychological Corporation)، لا يمكن الإستغناء عن دور عالم النفس المتخصص في توثيق تحليلات النتائج والإضافة عليها.

اختبار WAIS-III

تم في منتصف التسعينيات بناء وقياس نسخة معدلة عن اختبار WAIS-R قبيل نشرها عام ١٩٩٧ وذلك بعد أن أدى مرور الزمن إلى اعتبار محتوى ومعايير WAIS-R بالية وبحاجة إلى تجديد، وليس هذا بالأمر المستغرب نظراً لأن جميع اختبارات القدرة الإدراكية ومعايير ومحتوى اختبارات الذكاء تفقد قيمتها مع مرور الزمن. وقد تم الإعتناء عند إعادة النظر في اختبار WAIS-R بالاختبارات الفرعية المتعلقة بالمهارات اللغوية (كالمعلومات والمفردات والإستيعاب) وذلك لأن ارتباط هذا النوع من الاختبارات بالتغيرات الحضارية يجعلها عرضة لأن تصبح قديمة بالية أكثر من غيرها من الاختبارات الفرعية. ولم تقتصر التعديلات في اختبار WAIS-R على مراجعة الاختبارات الفرعية الأحد عشرة التي يتألف منها الاختبار بل تم كذلك إضافة ثلاثة اختبارات فرعية جديدة هي: الاستدلال البنيوي والبحث عن الرمز وسلاسل الأعداد والأرقام. ويتألف اختبار الاستدلال البنيوي من مجموعات من الصور لخمسة أشكال هندسية يطلب من المتحن التعرف على الشكل الصحيح منها عن طريق الإشارة أو التسمية، وتنتمي

فقرات هذا الاختبار الفرعي إلى الفئات التالية: إكمال النمط والتصنيف والاستدلال بالتشابه والاستدلال بالتسلسل. أما اختبار البحث عن الرموز فيتألف من عدد من المجموعات التي تم ترتيبها في أزواج يتألف كل زوج منها من مجموعة تضم الرمز المراد ومجموعة أخرى يبحث الممتحن فيها عن الرمز التي تضمه المجموعة الأولى ثم يضع إشارة في المربع الملحق بالسؤال تتناسب مع وجود الرمز أو عدم وجوده. ويعتمد الاختبار الفرعي لسلاسل الأرقام والأحرف على ترديد عدد من الأرقام والأحرف للممتحَن بشكل غير مرتب ثم يطلب منه إعادة ترديد ما ذكر مع ترتيب الأحرف أبجدياً والأرقام تسلسلياً من الأصغر إلى الأكبر.

وقد تم تقليص المدة الزمنية التي يستغرقها اختبار WAIS-III وخفض الدرجة الدنيا لمعظم الاختبارات الفرعية وذلك للحصول على تقييم أفضل للوظائف الأدراكية عند الأفراد المتخلفين عقلياً. ويساعدنا اختبار WAIS-III في الحصول على معدلات الذكاء IQs اللغوي التقليدي والأدائي والعام بالإضافة إلى درجات على مؤشرات العوامل الأربعة التالية: الإستيعاب اللغوي والذاكرة العملية والتنظيم الإدراكي – الحسي وسرعة معالجة المعلومات. أما بالنسبة للخصائص السايكومترية لاختبار WAIS-III فقد تم قياس الاختبار على عينة من ٢٤٥٠ فرداً بالغاً تراوحت أعمارهم بين ١٦ و٨٩ عاماً. وتم ترتيب العينة طبقياً تبعاً للعرق أو الأثنية (البيض والأفارقة الأمريكيين وذوي الأصول الإسبانية وغيرهم) ثم النوع (ذكر/أنثى) ثم المستوى التعليمي والموقع الجغرافي كلٌ ضمن فئته العمرية المناسبة. وتم معايرة اختبار WAIS-III مع النسخة المحسنة من مقياس ويكسلر للذاكرة (راجع الفصل الثامن من الكتاب) مما زودنا بمعلومات عن الفروقات في عمل الذاكرة وغيرها من القدرات الإدراكية عند الأفراد حتى سن ٨٩. ويضم دليل المستخدم للاختبار على نسب الارتباط بين درجات اختبار WAIS-III وغيره من الاختبارات في عائلة ويكسلر واختبار ستانفورد – بينيه

(الطبعة الرابعة) ومصفوفات ريفين التقدمية، كما يضم معلومات إحصائية عن مجموعات سريرية مختلفة.

اختبارات ديترويت لقابلية التعلم

تعد بطاريات اختبارات الذكاء الثلاثة التي صممها ريتشارد هاميل اختبارات مشابهة لمقاييس ويكسلر من حيث استهدافها لفئات عمرية مختلفة إلا أنها تختلف عن سلسلة ويكسلر في مفهومها وطرق إجرائها وتصحيحها. وتحمل الطبعات الحالية من هذه البطاريات الأسماء التالية:

- اختبار ديترويت لقابلية التعلم - ابتدائي (DLTA-P:2) وهو اختبار مصمم للأطفال بين ٣ إلى ١٠ سنوات؛
- اختبار ديترويت لقابلية التعلم - الطبعة الثالثة (DLTA-3) وهو اختبار مصمم للأطفال والمراهقين بين ٦ و ١٧ عاماً؛
- اختبار ديترويت لقابلية التعلم - البالغين (DLTA-A) وهو للبالغين بين ١٧ و ٧٩ عاماً.

يعد اختبار (DLTA-3) أكثر هذه الاختبارات شعبية إلا أنه لا مجال للمقارنة على الإطلاق بين استخدام هذه الاختبارات والاستخدام الواسع لاختبارات ويكسلر.

يتألف اختبار DLTA-3 من ١١ اختباراً فرعياً تقيس القدرات التالية:

١- الكلمات الأضاد (مفردات).

٢- سلاسل التصميم (التمييز البصري والذاكرة).

٣- تقليد الجمل (النحو).

٤- الأحرف المقلوبة (تذكر الترتيب والمهارات السمعية).

٥- إنشاء القصص (رواية القصص).

٦- إعادة رسم التصاميم (الرسم من الذاكرة).

٧- المعلومات الأساسية (معلومات عامة من الحياة اليومية).

٨- العلاقات الرمزية (الاستدلال البصري).

٩- سلاسل الكلمات (إعادة ترديد الكلمات).

١٠- سلاسل القصص (تنظيم الأجزاء ذات المعنى).

١١- أجزاء الصور (إتمام النواقص على مبدأ الغشتالت).

يستغرق اختبار DLTA-3 بين ٥٠ و١٢٠ دقيقة وتستند خطوات إجرائه إلى مزيج من الإجراءات المتبعة في اختبار ستانفورد - بينيه واختبارات ويكسلر، إذ يتم تحديد العمر القاعدي وسقف الاختبار لكل ممتحن ثم يمنح الممتحن درجة واحدة أو درجة الصفر على الأسئلة في الاختبارات الفرعية ١ إلى ٥ و ٧ إلى ٩ و١١ بينما يمنح الطالب ثلاث درجات أو درجة الصفر على أسئلة الاختبار الفرعي رقم ٦ و٥ درجات أو درجة الصفر على أسئلة الاختبار الفرعي رقم ١٠.

وكما كان الحال في اختبار ويكسلر، يتم تحويل الدرجات الأولية لاختبارات DLTA-3 الفرعية إلى درجات قياسية (بوسط حسابي = ١٠ وانحراف معياري = ٣) ثم يتم الجمع بين الدرجات القياسية للاختبارات الفرعية لنحصل على أربع أنواع من الدرجات المركبة (المحصّلة) هي: الدرجة المركبة العامة والدرجة المركبة للمستوى الأمثل والدرجات المركبة الميدانية والدرجات المركبة النظرية. يتم حساب الدرجة المركبة العامة والمبنية على مؤشر الدرجات القياسية ذو المتوسط

الحسابي ١٠٠ والإنحراف المعياري ١٥ عن طريق جمع الدرجات على المؤشر القياسي للاختبارات الفرعية الأحد عشرة، تعتبر الدرجة الناتجة عن هذه العملية من أفضل التقديرات لعامل الذكاء العام (g) عند الأفراد. تعد الدرجة المركبة للمستوى الأمثل (والتي يتم حسابها عن طريق جمع أعلى أربعة من الدرجات القياسية للاختبارات الفرعية الأحد عشرة) أكثر قدرة على تقدير الإمكانيات الإدراكية العامة لدى الممتحَن. ويمكن الحصول على الدرجات المركبة لستة مجالات معرفية (في الميادين اللغوية وغير اللغوية والانتباه المعزز والانتباه المقلّص والحركة المعززة والحركة المقلصة) عن طريق تلخيص الأداء العام للممتحَن في الاختبارات الفرعية المرتبطة بكل ميدان. أما بالنسبة إلى الدرجات المركبة النظرية فتستند هذه الدرجات إلى عدد من النظريات التي تتناول موضوع القدرات الإدراكية مثل: نظرية كاتتل وهورن حول الفرق بين الذكاء المتبلور والذكاء الفطري، ونظرية جينسين حول الفرق بين المستويات الإدراكية والمستويات الترابطية للتفكير، ونظرية داس حول الفرق بين المعالجة العقلية المتزامنة والمتتابعة للمعلومات، ونظرية ويكسلر حول الفرق بين القدرات الأدائية والقدرات اللغوية.

ومن الجدير بالذكر أنه وعلى الرغم من أن اختبار DLTA-3 يعد النسخة الأفضل من هذه البطارية إلا أن ثبات بعض اختباراته الفرعية لم يتم تحريه بالشكل الكافي، كما أن عينة قياس الاختبار لا تمثل المجتمع الذي تم إعداد الاختبار لأجله والمعلومات المتوفرة عن صدق الاختبار هزيلة للغاية.

بطارية كوفمان لتقييم الأطفال

تم تصميم بطارية كوفمان لتقييم الأطفال K-ABC كاختبار عادل حضارياً مؤلف من عدد من الاختبارات التحصيلية واختبارات الاستعداد التي تستهدف تقييم القدرات الإدراكية عند الأطفال بين أعمار سنتين ونصف إلى اثنتي عشرة

سنة ونصف. إن الجزء الخاص بالاستعداد في بطارية K-ABC بني على نظرية ألكسندر لوريا وج. ب. داس النفسية – العصبية التي تفيد بأن الذكاء يتألف من عدد من العمليات العقلية المتزامنة والمتتابعة. كما تفيد هذه النظرية بأن الفص الجداري والفص القذالي من القشر الدماغية هما اللذان يتحكمان بعملية المعالجة المتزامنة للمعلومات التي تسمح بالجمع والدمج بين المعلومات لتكوين وحدة متكاملة (ومن الأمثلة على هذه العملية إكمال الأجزاء الناقصة من صورة شبه مكتملة). أما المعالجة التتابعية (التسلسلية) للمعلومات (التي نحتاج إليها عند إستدعاء سلسلة من الأرقام) فيسيطر عليها الفصان الصدغي والجبهي للدماغ. وعلى الرغم من اقتناعه بوجود فروق بين المعالجة المتزامنة والمعالجة المتتابعة للمعلومات، فضل كوفمان ربط نوعي معالجة المعلومات بالقسم الأيمن والقسم الأيسر من الدماغ والذي يستند إلى نموذج روجر سبيري (١٩٦٨) الذي قسم الدماغ إلى فص أيمن وفص أيسر. إلا أنه لا يزال هنالك جدل كبير حول تحديد الموقع العصبي – النفسي لهذه العمليات وتحديد ما إذا كانت العمليتان تتطلبان مساهمة من الدماغ بأكمله وليس جزء أو فص معين منه.

يتألف اختبار K-ABC من ١٦ اختباراً فرعياً تستغرق بين ٣٠ و٨٠ دقيقة، وتتقسم الاختبارات التي تقيس عمليات معالجة المعلومات التتابعية إلى: ١) اختبار حركة اليدين. ٢) واختبار تذكر الأرقام. ٣) وترتيب الكلمات. أما الاختبارات الفرعية السبعة التي تقيِّم عملية معالجة المعلومات المتزامنة فهي: ١) النافذة السحرية، ٢) والتعرف على الوجوه، ٣) والإكمال على طريقة الغشتالت، ٤) والمثلثات، ٥) والتشابهات البنيوية، ٦) والذاكرة المكانية، ٧) وسلاسل الصور. وأخيراً، تتألف الاختبارات الفرعية التحصيلية الستة من: ١) المفردات المعبرة. ٢) الوجوه والأماكن، ٣) الحساب، ٤) الأحاجي، ٥) القراءة وفك الرموز، ٦) القراءة والفهم.

تعتمد طريقة إجراء الاختبار على تحديد نقطة البداية للممتحَن في كل اختبار فرعي والتي تستند إلى عمره الفعلي. ويستمر الممتحَن بالإجابة على الأسئلة حتى يصل إلى النقطة التي يتم عندها التوقف عن طرح أسئلة ضمن هذا الاختبار الفرعي. يتم بعد ذلك تحويل الدرجات التي حصل عليها الممتحن في الاختبارات الفرعية المتعلقة بالمعالجة المتزامنة والمعالجة المتتابعة للمعلومات والاختبارات الفرعية التحصيلية إلى مؤشر للدرجات القياسية بوسط حسابي يساوي ١٠ وإنحراف معياري يساوي ٣. كما يتم جمع درجات الاختبارات الفرعية المناسبة للحصول على أربعة معدلات درجات عامة في أربعة مجالات هي: معالجة المعلومات التتابعية، ومعالجة المعلومات المتزامنة، ومركّب المعالجة العقلية (معالجة المعلومات المتزامنة والمتتابعة)، والتحصيل.

تم قياس الاختبار على عينة من الأطفال اعتمدت على نتائج الإحصاء السكاني في الولايات المتحدة لعام ١٩٨٠، وقد تم ترتيب العينة طبقياً تبعاً للعمر والنوع والوضع الإجتماعي – الإقتصادي والفئة العرقية والمنطقة الجغرافية وحجم الجالية (التي يتنمي إليها أفراد العينة) والتوزيع التعليمي بالإضافة إلى مجموعة من الأطفال الموهوبين. ويحتوي دليل المستخدم الذي يرافق اختبار K-ABC على نتائج الأبحاث العديدة التي ارتبطت بالاختبار. وقد تم استخدام هذا الاختبار في المؤسسات التعليمية والعيادات النفسية لأغراض نفسية – تعليمية قادت إلى التعرف على الأطفال المتأخرين عقلياً والأطفال الموهوبين على حدٍ سواء. كما ساهم الاختبار في عمليات توزيع الأطفال على الصفوف التعليمية المختلفة وفي عملية التخطيط التعليمي وفي تقييم أفراد الأقليات العرقية في المجتمع والأطفال دون سن المدرسة بالإضافة إلى التقييم العصبي – النفسي ومجالات البحث العلمي.

يحدد كتاب "دليل المستخدم في تصحيح اختبار K-ABC وتفسير نتائجه" (كوفمان وكوفمان، ١٩٨٣a) الخطوات التي يجب اتباعها لتحليل نتائج هذا

الاختبار والتي تنطوي على مقارنة درجات الممتحَن في المؤشرات العامة للاختبار وتحديد نقاط قوته وضعفه في الاختبارات الفرعية المرتبطة بالمعالجة العقلية للمعلومات وبالتحصيل. ويساعد إعداد ملف للدرجات القياسية لكل اختبار فرعي وللمؤشرات العامة للاختبار في عملية تفسير نتائج اختبار K-ABC، كما يحتوي كتاب "دليل المستخدم لتفسير نتائج K-ABC" (كوفمان وكوفمان، b١٩٨٣) على ملفات لمعدلات درجات أطفال عندهم صعوبات تعلم وأطفال متأخرين عقلياً وأطفال عندهم مشاكل سلوكية وأطفال معاقون جسدياً وأطفال الخطر العالي (أطفال ما دون سن المدرسة الذي لم تتضح إمكانيتهم بعد) وأطفال عندهم صعوبات في السمع والأطفال المميزون والموهبون.

لقد نجح اختبار K-ABC لدرجة كبيرة في تحقيق هدف بناء اختبار عادل حضارياً لقياس القدرات الإدراكية. ومع ذلك فقد تم انتقاد هذا الاختبار لمجموعة من النقاط أهمها:

- عدم وجود تطابق بين الاختبار وبين الأساس النظري الذي استند إليه.
- الخلط بين استخدام مصطلحي "المعالجة العقلية" و"التحصيل" في تسمية الاختبارات الفرعية.
- ازدياد التركيز على المعالجة التتابعية للمعلومات على حساب المعالجة المتزامنة وذلك في اختبار المعالجة العقلية المركّب.
- التركيز الشديد على المهام التي تعتمد على الذاكرة قصيرة الأمد وعلى حدة الانتباه.
- الإبهام وعدم الوضوح في استخدام المعالجة المتزامنة والمعالجة المتتابعة (التسلسلية).
- التصنيف غير الدقيق لحالات التخلف العقلي في مختلف الفئات العمرية للمقياس.

- انخفاض سقف الاختبار مما يحد من استخداماته مع الأطفال الموهوبين؛
- وجود مشاكل كثيرة وكبيرة في عينة القياس؛
- وجود صعوبات في تقييم فعالية الخطط التدريسية المبنية على ملفات الأطفال الناتجة عن استخدام اختبار K-ABC (ساقلر، ١٩٨٨).

اختبارات K-BIT وKAIT

صمم آلان كوفمان اختبارين آخرين يستخدمان على نطاق أضيق بكثير من استخدام اختبار K-ABC هما: اختبار كوفمان المختصر للذكاء (K-BIT) واختبار كوفمان لذكاء المراهقين والبالغين (KAIT)، ويستند الاختباران إلى نظريات ر. ب. كاتل حول الذكاء المتبلور والذكاء الفطري[٢] ويغطيان مدى عمري واسع هو ٤ إلى ٩٠ عاماً في اختبار K-BIT و١١ إلى ٨٥+ في اختبار KAIT. ويستغرق اختبار KAIT بين ٦٠ إلى ٩٠ دقيقة بينما لا يحتاج اختبار K-BIT إلى أكثر من ١٥ إلى ٣٠ دقيقة.

يتألف اختبار K-BIT من اختبارين فرعيين هما اختبار للمفردات يقيس القدرة اللغوية أو القدرة المتبلورة واختبار للبنى يقيس القدرات اللالغوية أو القدرات الفطرية، ويمكن الحصول على درجات الاختبارين الفرعيين بالإضافة إلى معدل ذكاء مركّب IQ يتم حسابه عن طريق جمع الدرجات القياسية على المؤشر لكلا الاختبارين.

تضم بطارية KAIT الأساسية ثلاثة اختبارات فرعية تقيس القدرة المتبلورة (إستيعاب المسموع والمعاني المزدوجة والتعريفات) وثلاثة اختبارات فرعية تقيس

---

(٢) تولد القدرة الفطرية Fluid ability مع الفرد وتلعب الجينات فيها دوراً حاسماً ونراها غالباً في مواقف حل المشكلات والاستجابات الإبداعية للمحرضات الخارجية. أما القدرة المتبلورة فهي نتاج للمعرفة والمهارات التي يكتسبها المرء من خبراته وتعليمه.

القدرة الفطرية (تعلم الرموز اللغوية والرموز الغامضة والخطوات المنطقية)، أما بطارية KAIT الموسعة فتضم بالإضافة إلى الاختبارات الفرعية الستة هذه أربعة اختبارات أخرى هي: تذكر تصاميم المكعبات ووجوه مشهورة وتذكر الكلمات المرمزة وتذكر المسموع. يتم تحويل الدرجات الأولية في الاختبارات الفرعية العشرة إلى درجات قياسية وسطها الحسابي ١٠ وإنحرافها المعياري ٣. ثم تجمع الدرجات المعيارية للاختبارات الفرعية لتعطينا معدل الذكاء الفطري والمتبلور والمركّب على مقياس وسطه الحسابي ١٠٠ وإنحرافه المعياري ١٥.

اختبارات سلوسون للذكاء

يعد اختبار سلوسون للذكاء واحداً من اختبارات الذكاء الفردية المتعددة المستويات التي تتمتع بقاعدة شعبية لا بأس بها، وتتوفر حالياً النسخ التالية من هذا الاختبار:

- اختبار سلوسون للذكاء (المرحلة الإبتدائية) (SIT-P) للأطفال بين ٢ إلى ٧,١١ سنة؛
- اختبار سلوسون المعدل للذكاء (SIT-R) للأفراد بين ٤ و٦٥ عاماً؛
- اختبار سلوسون الشامل (S-FRIT) للأفراد من ٥ سنوات إلى سن البلوغ.

وتشابه هذه الاختبارات اختبار K-BIT من حيث كونها اختبارات لا يستغرق إجراؤها وتصحيحها أكثر من ١٥ إلى ٣٠ دقيقة. وقد تم قياس اختبار SIT-R عام ١٩٩٠ وهو اختبار مؤلف من ١٨٧ سؤالاً تختبر المفردات والمعلومات العامة ووجوه التشابه والإختلاف والإستيعاب والقدرة الكمية والذاكرة السمعية. وتُحسب الدرجات الأولية عن طريق إضافة معدل الدرجات القاعدي إلى عدد الفقرات التي تمت الإجابة عليها بشكل صحيح بعد تحديد المستوى القاعدي. ويتم التعبير عن الأداء العام للأفراد في هذا الاختبار على شكل معدل

ذكاء انحرافي deviation IQs أو معدل درجات قياسي كلى متوسطه الحسابي ١٠٠ وإنحرافه المعياري ١٦.

مقياس القدرة التباينية

لم يتعرض مقياس القدرة التباينية (DAS) لنفس الدرجة من الإنتقاد التي تعرض لها اختبار SIT-R بسبب عملية قياسه التي لم تتم بالشكل المطلوب. واختبار DAS هو النسخة الأمريكية المعدلة عن اختبار BAS أو مقياس القدرة البريطاني. صُمم اختبار DAS للأطفال بين عامين ونصف إلى ١٧ عاماً ويتألف الاختبار من ١٩ اختباراً فرعياً للقدرة تم تقسيمها إلى أربع نواحٍ أساسية للقدرة هي: القدرة اللغوية واللالغوية والمكانية والتشخيصية. ولا يمكن إجراء أكثر من ثمانية إلى إثنى عشرة اختباراً لمستوى عمري معين. وتستغرق الاختبارات الفرعية كافة ما لا يزيد عن ٤٥ إلى ٦٥ دقيقة. وكما كان الحال في اختبار K-ABC يمكن إلحاق عدد من اختبارات التحصيل (قراءة الكلمات والتهجئة والمهارات العددية الأساسية) باختبار BAS.

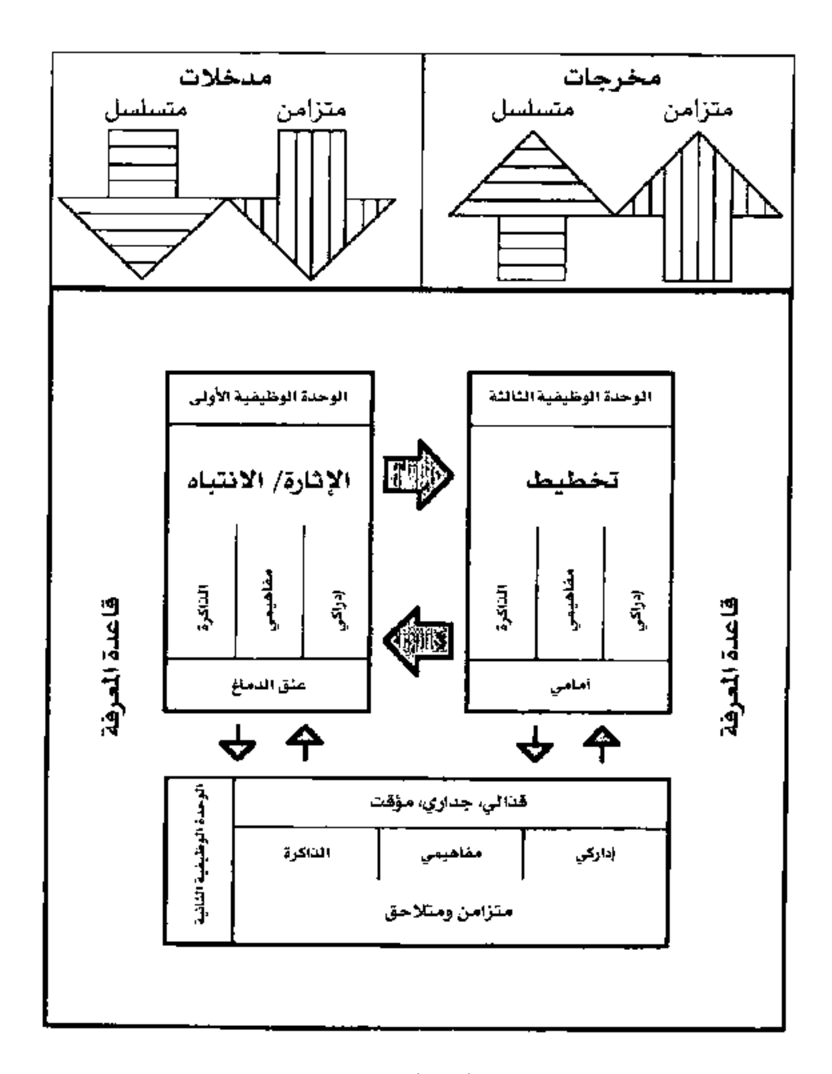
يتم تحويل الدرجات الأولية في اختبارات BAS الفرعية إلى درجات قدرة ودرجات T ورتب مئوية مبنية على عينة القياس لعام ١٩٨٨، كما يمكن الجمع بين درجات الاختبارات الفرعية للحصول على درجات كلية للقدرة اللغوية وقدرة الإستدلال غير اللغوي والقدرة المكانية والقدرة التصورية العامة. ويتم التعبير عن هذه الدرجات على مؤشر الدرجات القياسي متوسطه الحسابي ١٠٠ وانحرافه المعياري ١٥، وعلى حين يتم التعبير عن درجات الاختبارات الفرعية التشخيصية التي تقيس المهارات الإدراكية – الحسية ومهارات الذاكرة عن طريق درجات T ودرجات مئينية، يتم التعبير عن درجات الاختبارات الفرعية التحصيلية على شكل مكافئات درجة qrade equivalents ورتب مئينية ودرجات قياسية عمرية. ونظراً لأنه تم الحصول على معايير الاختبارات الفرعية التي تقيس القدرة وتلك

التي تقيس التحصيل من نفس عينة الأشخاص، يمكن إجراء مقارنات مباشرة وتشخيصية بين قدرات الممتحَن التي تم قياسها وتحصيله الفعلي.

تتميز اختبارات BAS ولا سيما تلك التي تعنى بالقدرة اللغوية والإستدلال اللالغوي والقدرة المكانية والقدرة التصورية العامة بنسبة ثبات عالية، كما أنه لم يتم اكتشاف أي أثر للتحيز العرقي في نتائج أداة القياس هذه.

نظام داس – ناغلييري للتقييم الإدراكي (CAS)

يعد هذا النظام شأنه كشأن اختبار K-ABC اختباراً جديداً للذكاء يستند إلى أسس نظرية آ. ر. لوريا العصبية – النفسية. تم تصميم اختبار CAS بناءً على نموذج PASS (التخطيط والانتباه والمعالجة المتزامنة والمعالجة المتلاحقة) لعمل الدماغ الذي يقسم وظائف الدماغ إلى ثلاث وحدات (انظر الشكل ٧–٢). تُعنى الوحدة الوظيفية الأولى التي ترتبط بجذع الدماغ وبالجهاز الحوفي (اللمبي) Limbic system بعمليات الإثارة والانتباه والتمييز بين المحفزات المختلفة، ولا تتدخل هذه الوحدة بشكل مباشر في عملية استقبال وتحليل المعلومات بل في العمليات التي تسبقها والتي تشمل الاستعداد (الجاهزية) لاستقبال المعلومات وتركيز الانتباه. وتركز الوحدة الوظيفية الثانية التي تضم مراكز البصر والسمع والإحساس العام في فصوص الدماغ القذالية والصدغية والجدارية على استقبال المعلومات ثم تحليلها أو الإضافة عليها وتنميقها ثم تخزينها عن طريق عمليات المعالجة المتزامنة والمتلاحقة. وتقوم الوحدة الوظيفية الثالثة والتي تقع في المنطقة الأمامية من نصفي الكرة المخية بعمليات البرمجة للعمليات الإدراكية وتنظيمها والتحقق من عملها، كما أنها تنظم عمل الوحدة الوظيفية الأولى بحيث تسمح بتوافق أفعال الشخص مع نواياه وأهدافه الواعية.



الشكل ٧-٢

نموذج PASS للقدرات الإدراكية (من داس وناغلييري وكيربي، ١٩٩٤، ص٢١، حقوق الطبع ١٩٩٤ لآلين وبيكون. تمت إعادة الطبع بناءً على الموافقة)

وكما يوضح الشكل ٧-٢، تستند الوحدات الوظيفية الثلاثة إلى القاعدة المعرفية للفرد والتي تضم كافة المعلومات المخزنة في ذاكرة الفرد والتي يمكن استدعاؤها لتدخل في عملية معالجة المعلومات الجديدة. وتتحد القاعدة المعرفية

للفرد مع مركز التخطيط (الوحدة الوظيفية الثالثة) أومركز الإثارة والانتباه (الوحدة الوظيفية الأولى) أو مركز المعالجة المتزامنة والمتلاحقة (الوحدة الوظيفية الثانية) تبعاً لمقتضيات المهمة المراد إنجازها. ويوضح الشكل ٧-٢ كذلك أنه يمكن لكل من المدخلات الحسية والمخرجات الحركية (التحدث أو الكتابة...إلخ) أن تكون متسلسلة أو متزامنة (متلازمة).

يتألف اختبار CAS المصمم للأطفال بين ٥ إلى ١٧ عاماً و١١ شهراً من ثمانية اختبارات فرعية في البطارية الأساسية و١٢ اختباراً فرعياً في البطارية النموذجية. وكما يوضح الجدول ٧-٤ يضم CAS ١٢ اختباراً فرعياً إلا أنه لا يمكن إجراء أكثر من ٨ إلى ١٢ اختباراً للفرد الواحد.

الجدول ٧-٤

اختبارات نظام التقدير الإدراكي الفرعية

التخطيط

- مطابقة الأرقام: يجد المتحَن استراتيجية تمكنه من إيجاد رقمين متشابهين في عدد من السطور.
- الرموز المخطط لها: يستخدم المتحَن خطة لملء الرموز المطابقة للأحرف بأسرع وقت ممكن.
- العلاقات المخطط لها: يستخدم المتحَن طريقة ما تمكنه من إتمام نمط مؤلف من أحرف وأرقام.

الانتباه

- الانتباه التعبيري: يسمي المتحَن اللون الذي كتبت به كلمة تشير إلى أحد الألوان على الرغم من عدم التطابق بين لون الطباعة واللون الذي تمثله الكلمة.
- التحري عن الأرقام: يتعرف المتحن على أرقام محددة تتطابق مع العينة المعطاة ويقاوم الحيرة في أمر الأرقام المموهة التي تهدف إلى تضليله وجذب انتباهه.

- الانتباه الإستقبالي: يتعرف المتحن على أزواج الصور والأحرف الصحيحة ولا يعير اهتماماً للمموهات.

المعالجة المتزامنة

- البنى اللالغوية: يختار المتحن واحداً من ستة خيارات لإتمام بنية لا لغوية متدرجة.
- العلاقات اللغوية المكانية: يختار المتحن واحدة من ست صور تجيب على سؤال متعلق بالعلاقات المكانية.
- تذكر الأشكال: يتعرف المتحن على شكل هندسي معين تم وضعه ضمن تصميم معقد.

المعالجة المتلاحقة (التتابعية)

- سلاسل الكلمات: يكرر المتحن مجموعة من الكلمات بالترتيب بعد فراغ القائم على الاختبار من قراءتها بصوتٍ عالٍ.
- تكرار الجمل: يكرر المتحن جملاً لها قوام نحوي إلا أن معانيها غيرمتماسكة.
- معدل النطق (لأعمار ٥-٧ سنوات): يكرر المتحن مجموعة من الكلمات ١٩ مرة بسرعة وبدون أخطاء.
- أسئلة الجمل (لأعمار ٨-١٧ عاماً): يجيب المتحن على أسئلة عن جمل معينة بالاعتماد على الصيغة القواعدية للجملة.

- متضمن في البطاري الأساسية

المصدر: أعيد الطبع بتصرف مع موافقة شركة ريفرسايد للنشر، ١٩٩٧.

تناسب الاختبارات الفرعية في بطارية CAS الأطفال من مختلف الأعمار باستثناء اختبار معدل النطق والذي يناسب الأطفال بين أعمار الخامسة والسابعة واختبار أسئلة الجمل المعد للأطفال بين أعمار ثمانية وسبعة عشرة. تستغرق البطارية الأساسية ٤٠ دقيقة بينما تستغرق البطارية القياسية ٦٠ دقيقة.

تدخل ثلاثة متغيرات هي عدد الإجابات الصحيحة والزمن والإكتشافات (التحريات) الكاذبة في حساب درجات اختبارات CAS الفرعية، وتُجمع درجات

الفرد في الاختبارات الفرعية الثمانية أو الإثنى عشرة للحصول على أربع درجات PASS قياسية (ترمز PASS إلى عمليات التخطيط والانتباه والمعالجة المتزامنة والمتتابعة التي تحدثنا عنها سابقاً)، وتُجمع هذه الدرجات بدورها لتعطينا درجات المقياس الكامل. ويوفر دليل المستخدم لبطارية CAS معلومات عن تحويل الدرجات الأولية في المقاييس الأربعة وفي المقياس الكامل إلى درجات قياسية ورتب مئينية ومكافئات عمرية (فواصلها مؤلفة من ٤ أشهر).

تم قياس اختبارات CAS في منتصف التسعينيات من القرن الماضي على عينة تمثل المجتمع من الأطفال والمراهقين في الولايات المتحدة. ويضم "دليل تفسير النتائج" معاملات ثبات الإتساق الداخلي ومعاملات إعادة الاختبار لبطارية CAS والتي تقع ضمن ٩٠٪ لدرجات المقياس الكامل وضمن ٨٠٪ إلى أول ٩٠٪ لمقاييس PASS الأربعة. ويحتوي دليل مستخدم CAS على نتائج مجموعة من الدراسات متعلقة بصدق الاختبار تم إجراؤها على أطفال متخلفين عقلياً وأطفال يعانون من صعوبات في التعلم بالإضافة إلى الأطفال الموهوبين وأطفال يعانون من نقص الانتباه أو من إصابات دماغية. تعد بطارية CAS من أدوات القياس الواعدة والمثيرة للاهتمام إلا أننا لا نزال بحاجة إلى معلومات إضافية عن الخصائص السايكومترية لهذه البطارية وعن الفائدة التي يمكن الحصول عليها منها عند إتخاذ قرارات حول الاحتياجات العلاجية أوالتعليمية للأطفال.

اختبارات الذكاء الجماعية

يعود الفضل في وضع أول اختبار ذكاء جماعي لآرثر أوتيس الذي عدّل اختبار ستانفورد - بينيه وتوصل إلى اختبار ورقة وقلم لتقييم الذكاء عند المجموعات وليس الأفراد فقط. وقد قادت جهود أوتيس في هذا المجال إلى تطوير أول اختبار جماعي قياسي للذكاء وهو اختبار الجيش ألفا ونظيره

للمجندين الأميين، اختبار الجيش بيتا. وقد ساهم هذان الاختباران اللذان تم إجراؤهما للآلاف من الجنود الأمريكيين خلال وبعد الحرب العالمية الأولى في توفير نموذج عن اختبارات الذكاء الجماعية أدى إلى ظهور الاختبارات التالية: مقياس أوتيس لذكاء المجموعات، والمقياس الوطني للذكاء، واختبارات أوتيس الذاتية للقدرة الذهنية، واختبارات كولمان – أندرسون للذكاء، واختبارات هينمون – نلسون للقدرة العقلية. ويمكن للقارئ الحصول على قائمة مختارة لاختبارات الذكاء المتوفرة حالياً في الأسواق عن طريق اختيار البند١ من قائمة برنامج H-7 الموجود على القرص المرفق بالكتاب.

الصيغة والوظائف

على الرغم من احتواء اختبارات الذكاء الجماعية الأولى على أسئلة تتطلب اجابات قصيرة، مما لا شك فيه أن أسئلة الخيارات المتعددة هو الشكل الأكثر استخداماً في اختبارات الذكاء. تُرتب الفقرات الاختبارية في اختبارات الذكاء الجماعية عادةً إما على شكل كتيب امتحاني لولبي تزداد صعوبة أسئلته كلما تقدم الممتحَن في الاختبار (كاختبار أوتيس – لينون للقدرة المدرسية واختبارات هينمون – نلسون للقدرة العقلية) أو على شكل اختبارات فرعية منفصلة لها زمنها المحدد الخاص بها (كاختبار المهارات الإدراكية واختبار القدرات الإدراكية). وتختلف اختبارات الذكاء الجماعية بالإضافة إلى اختلافها في طريقة ترتيب فقراتها في المدى العمري للاختبار أو في عدد المستويات العمرية التي يصلح استخدام الاختبار معها. وتضم معظم الاختبارات المرموقة عدداً من المستويات بحيث توفر صور متعددة للاختبار أو مجموعات متعددة من الأسئلة تتناسب مع الفئة العمرية للممتحَنين.

وقد تم استخدام اختبارات الذكاء الجماعية كمرحلة اختيار أولية في المجالات التعليمية والتوظيفية وذلك لتحقيق ما يلي:

- التعرف على الطلاب الذين يحتاجون إلى اختبار ذكاء فردي.
- اختيار المتقدمين إلى وظيفة معينة.
- اختيار الطلاب لإلحاقهم ببرنامج تعليمي أو تدريبي معين.
- اختيار الطلاب الذين سيقبلون في جامعة أو كلية أو أي مؤسسة تعليمية أخرى.

ويمكن استخدام اختبارات الذكاء الجماعية كجزء من بطارية من الاختبارات والاستبانات المستخدمة في مجالات الإرشاد التعليمي والمهني والذاتي، كما تستخدم اختبارات الذكاء الجماعية بشكل كبير في الأبحاث العلمية في مجالات التعليم والصناعة والجيش والدولة وغيرها من المنظمات والمؤسسات.

الخصائص السايكومترية

تنخفض معاملات ثبات إعادة الاختبارلاختبارات الذكاء الجماعية التي يعاد إجراؤها في فترات زمنية متباعدة عن معاملات الثبات لاختبارات الذكاء الفردية، وغالباً ما تميل معاملات الثبات لتكون أعلى في اختبارات الذكاء الجماعية اللغوية من نظائرها اللالغوية (هوبكينز وبراخت، ١٩٧٥). وتتميز اختبارات الذكاء الجماعية بأن لها نسبة ارتباط مرتفعة مع بعضها ومع درجات الأفراد في اختبارات الذكاء الجماعية. وتتراوح معدلات ارتباط درجات اختبار الذكاء الجماعية مع الدرجات المدرسية بين ٠,٤٠ إلى ٠,٧٠ وترتفع نسبة الإرتباط بين درجات الاختبار والمتغير الذي يهدف الاختبار إلى التنبؤ به مع إزدياد التشابه بين فقرات اختبار الذكاء الجماعي ومتغير المحك موضع القياس. وهكذا نلاحظ ارتفاع نسبة الترابط بين الاختبارات الجماعية اللغوية ومحكات التقييم اللغوية وبين الاختبارات الجماعية الكمية ومحكات التقييم الكمية.

اختبار أوتيس - لينون للقدرة المدرسية

ينحدر اختبار أوتيس - لينون للقدرة المدرسية (OLSAT) مباشرة من أول اختبار ذكاء جماعي وضعه آرثر أوتيس، وقد وصل هذا الاختبار إلى الطبعة السابعة (OLSAT7) وهو اختبار ذو مدى واسع يقيم قدرات الطلاب بين ٥ و١٨ عاماً أي من سن الحضانة إلى التخرج من المدرسة. ويضم الاختبار مستويات منفصلة للأطفال من الحضانة إلى الصف الثالث وذلك بسبب التطورات الذهنية والثقافية الهائلة التي تطرأ على الأطفال في هذا العمر. تتغير المدة الزمنية التي يستغرقها الاختبار بتغير مستوى الاختبار، إلا أن الوقت الأقصى للاختبار هو ٧٥ دقيقة.

يضم اختبار OLSAT7 أحد وعشرين اختباراً فرعياً تنقسم إلى فئتين: الاختبارات اللغوية (وتضم اختبارات الإستيعاب اللغوي والاستدلال اللغوي) والاختبارات اللالغوية (وتضم اختبارات الإستدلال عن طريق الصور والإستدلال عن طريق الأشكال والإستدلال الكمي). ويستند الأساس لبناء فقرات اختبارات OLSAT7 إلى فرضية أن تعلم الأشياء الجديدة يتطلب إدراكاً - حسياً صحيحاً، والتعرف على وتذكرما تم إدراكه حسياً، ثم التفكير المنطقي، وفهم العلاقات، وإستنباط مجموعة من الخصائص والتفاصيل، ومن ثم تعميم ما تم التوصل إليه على المواقف الجديدة.

تمت مطابقة OLSAT7 مع سلسلة اختبارات ستانفورد التحصيلية (الطبعة التاسعة) مما سمح بمقارنة قدرات الطالب مع تحصيله المدرسي، ويتم التعبير عن نتائج OLSAT7 على شكل مؤشر للقدرة المدرسية (مقسم حسب العمر) أو على شكل رتب مئينية أو رتب من تسعة STANINE[٣] أو المكافئ للمنحنى الطبيعي NCE (بناءً على العمر والصف الدراسي).

---

(٣) وهي طريقة لحساب درجات الاختبارات على مقياس معياري مؤلف من تسع نقاط وسطه الحسابي ٥ وإنحرافه المعياري ٢، (المترجم).

اختبار وندرليك للموظفين

يعد هذا الاختبار والذي تم تطويره عن اختبار أوتيس الذاتي للقدرة العقلية من أقصر اختبارات الذكاء الجماعية وأكثرها شعبيةً. يتألف هذا الاختبار من ٥٠ فقرة من نوع الخيارات المتعددة تستغرق ١٥ دقيقة فقط (انظر الشكل ٧-٣) وتضم متشابهات وتعريفات ومسائل للتفكير المنطقي ومسائل حسابية ومسائل عن العلاقات المكانية ومقارنة بين الكلمات وتحديد وجهة السير والإتجاهات.

تم استخدام اختبار وندرليك وبنسبة مقبولة من النجاح في مجال اختيار الموظفين في مجالات الصناعة والأعمال، وتعتبر نسب ثبات الاختبار وارتباط درجاته مع مقاييس الذكاء الأخرى نسب جيدة ولا سيما في حالة اختبار قصير كهذا. إلا أن احتمالية إساءة استخدام الدرجات الفاصلة للاختبار من قبل مدراء غير مدربين يثير بعض التحفظات على الاختبار (شميدت وشونفلدت ١٩٨٥).

الشكل ٧-٣

عينة من اختبار وندرليك للموظفين. أعيد الطبع بموافقة شركة اختبار وندرليك للموظفين المساهمة، ليبرتيفيل

١- انظر إلى مجموعة الأرقام أدناه وحدد الرقم الذي يكمل النمط:

٨ ٤ ٢ ١ ١/٢ ١/٤

٢- افترض صحة الجملتين الأولى والثانية. هل الجملة الثالثة:

أ- صحيحة ب- خاطئة جـ- لا يمكن الجزم

يمارس الولد لعبة البيسبول. يرتدي جميع لاعبي البيسبول قبعة.

الولد يرتدي قبعة.

٣- تكلف رزمة أوراق ٢١ سنتاً. ما كلفة ٤ رزم؟

٤- ما عدد الأزواج المتطابقة تماماً مما يلي:

نيمان، ك. م.	نييمان، ك. م.
توماس، ج. ك.	توماس، س. ك.
هوف، ج. ب.	هوف، ج. ب.
بينو، ل. ر.	بينوا، ل. ر.
وارتر، ت. س.	وارتر، ت. س.

٥- يستاء، يحجز

هل الكلمتان أعلاه: أ- مترادفات ب- متعاكسات ج- لا هذه ولا تلك

٦- يختلف أحد الأشكال التالية عن الأشكال الباقية بشكل كبير. ما هو هذا الشكل؟

1 2 3 4 5

٧- يقطع قطار مسافة ٢٠ قدم في ١/٥ من الثانية. ما المسافة التي يقطعها القطار في ٣ ثواني؟

٨- إذا كان الحبل يكلف ١٠ سنتات للقدم، كم قدم يمكنك أن تشتري ب٦٠ سنتاً؟

٩- الشهر التاسع في السنة هو:

أ- أكتوبر ب- يناير ج- يونيو د- سبتمبر هـ- مايو

١٠- أي من الأرقام التالية له القيمة الأدنى:

٧ ٨ ٣١ ٣٣ ٢

١١- يرغب ناشر في نشر مقال من ٤٨٠٠٠ كلمة ويريد أن يستخدم حجمين من الطباعة. تضم الصفحة عند استخدام الحجم الكبير ١٨٠٠ كلمة و٢٤٠٠ كلمة عند استخدام الحجم الصغير. ومع العلم بأن عدد الصفحات المحدد للمقال في إحدى المجلات هو ٢١، ما هو عدد الصفحات التي يجب أن يستعمل الناشر فيها الحجم الصغير؟

١٢- يتشابه توزيع ساعات الليل والنهار في شهر سبتمبر مع شهر:

أ- يونيو ب- مارس ج- مايو د- نوفمبر

١٣- يشترك ثلاثة رجال في صفقة ويقررون توزيع الأرباح بالتساوي. يساهم (أ) ب ٩٠٠٠ دولار و(ب) ب ٧٠٠٠ دولار و(ج) ب ٤٠٠٠ دولار. إذا ربح الشركاء ٤٨٠٠ دولار، كم ينقص (أ) عن المبلغ الذي كان من الممكن أن يحصل عليه في حال قرر الشركاء تقسيم الأرباح حسب نسبة المساهمة؟

١٤- افترض صحة الجملة الأولى والجملة الثانية. هل الجملة الثالثة:

أ- صحيحة ب- خاطئة ج- لا يمكن الجزم

ألقى توم التحية على بيث. بيث ألقت التحية على دونا.

توم لم يلق التحية على دونا.

١٥- عمر شاب ما ١٧ عاماً وعمر أخته هو ضعف عمره. كم سيكون عمر الأخت عندما يكون عمر الأخ ٢٣ عاماً؟

هذه عينة من أسئلة الاختبار هدفها العرض فقط. شركة اختبار وندرليك للموظفين المساهمة هي الناشر الرسمي لاختبار وندرليك للموظفين.

الإجابات:

١ - ٨/١	٢- صحيحة	٣- ٨٤ سنتاً
٤- ١	٥- ٣	٦- ٤
٧- ٣٠٠ قدم	٨- ٦ أقدام	٩- سبتمبر
١٠- ٠,٣٣	١١- ١٧	١٢- مارس
١٣- ٥٦٠ دولاراً	١٤- لا يمكن الجزم	١٥- ٤٠ عاماً

لمحة عن بعض اختبارات الاستعداد الأكاديمي متعددة المستويات

تعد الاختبارات الثلاثة التالية من أكثر بطاريات الذكاء الجماعية المتعددة المستويات استخداماً في الصفوف المدرسية المتعددة ولسنوات طويلة:

- اختبار كاليفورنيا المختصر للنضج العقلي؛
- الاختبار المختصر للاستعداد الأكاديمي؛

- اختبار لورج – ثورندايك للذكاء.

وقد تم استبدال الاختبارين الأول والثاني من القائمة أعلاه باختبار المهارات الإدراكية (TCS) والاختبار الثالث باختبار القدرات الإدراكية (CogAT). ويمكن استخدام TCS وCogAT مع الصفوف المدرسية المختلفة (من الحضانة إلى الثانوية العامة) ويستغرق اختبار CogAT الذي يتألف من عدد أكبر من الاختبارات الفرعية وينقسم إلى عدد أكبر من المستويات ٩٠ إلى ٩٨ دقيقة على حين لا يستغرق اختبار TCS أكثر من ٥٠ إلى ٥٤ دقيقة. تقسم البطاريات اللغوية واللالغوية والكمية إلى اختبارين فرعيين لكل بطارية على مستويين ١ و ٢ في اختبار TCS وإلى ثلاثة اختبارات فرعية على مستويات من A إلى H في اختبار CogAT، وتتألف المستويات الستة لاختبار TCS من أربعة اختبارات فرعية (السلاسل والمتشابهات والذاكرة والإستدلال اللغوي) تعطينا ثلاثة مقاييس للاستعداد الأكاديمي هي المقياس اللغوي والمقياس اللالغوي ومقياس الذاكرة. ومن الجدير بالذكر أن هنالك اختبار خاص في بطارية TCS للأطفال في مرحلة الحضانة والصف الأول. تم قياس TCS وCogAT بطريقة تسمح بمقارنة الاستعداد الأكاديمي للطلاب مع درجاتهم في الاختبارات التحصيلية القياسية. وبشكل عام يمكن القول بأن TCS وCogAT اختباران يتمتعان بدرجة عالية من البنية، والاحترام والثبات كمقاييس للقدرة الإدراكية. كما أن الشركتين الموزعتين لهذين الاختبارين هما شركتان لهما سمعة طيبة وهما شركة ريفرسايد للنشر (CogAT) وماكغروهيل/CTB (TCS).

الاختبارات اللالغوية والاختبارات العادلة حضارياً

نظراً لأن القدرة على الكتابة والقراءة والتحدث هي المحك الأساسي للنجاح المدرسي، ليس من المستغرب أن الأطفال الأكثر تميزاً في المجال اللغوي هم الذين يحققون أفضل النتائج في التحصيل الأكاديمي. ومن الضروري أن نعرف أن

التدريب اللغوي الذي قد يحصل عليه بعض الأطفال قد لا يحصل عليه أطفال غيرهم كما أن بعض الأطفال قد لا يستفيدون بالشكل الكافي من التدريب اللغوي الذي يتلقونه، وقد لوحظ بناءً على هذا أن نتائج الأفراد في اختبارات الذكاء ذات التركيز اللغوي الكبير مثل اختبار ستانفورد - بينيه واختبار الجيش ألفا قد لا تكون مؤشراً دقيقاً عن الإمكانيات العقلية لهؤلاء الأفراد. وحتى وإن كان السبب وراء فشل بعض الأفراد في تحقيق نتائج جيدة في الاختبارات ذات التركيز اللغوي هو نقص في الموهبة وليس نقص في الخبرة التعليمية، يجب أن نأخذ بعين الاعتبار أن فقدان الموهبة اللغوية لا ينم عن ضعف عام في القدرات العقلية.

لم يساعد اختبار الجيش ألفا والذي تم إجراؤه لأعداد كبيرة من الشباب الأمريكيين خلال الحرب العالمية الأولى في الكشف عن القدرات الإدراكية الحقيقية للممتحَنين وذلك لأن التحصيل العلمي لمعظم الشباب آنذاك لم يكن يتجاوز المرحلة الإبتدائية. وقد أدى هذا إلى تطوير اختبار الجيش بيتا وهو اختبار لالغوي يضم فقرات تتعلق بتحليل المكعبات والرموز العددية والمجسمات الهندسية والمتاهات وإكمال الصور وغيرها. وقد تم تطوير عدد كبير من الاختبارات المشابهة إما كجزء من بطارية اختبار تضم أقسام لغوية ولالغوية (مثل اختبار القدرات الإدراكية) أو كأداة قياس مستقلة كاختبار الذكاء اللالغوي واختبار غودينوف - هاريس للرسم (وهما اختباران للذكاء الفردي)، ومقياس القدرة العامة للبالغين، واختبار البنى المتشابهة، واختبار ناغلييري اللالغوي للقدرة (وهي اختبارات يمكن إجراؤها للعديد من الطلاب في وقتٍ واحد) (انظر الشكل ٧-٤). وتستخدم كافة اختبارات الذكاء اللالغوية الفردية والجماعية في تقييم قدرات الأفراد الذين لا تعتبر اللغة الإنجليزية لغتهم الأم والأميين أو شبه الأميين أو الأفراد ذوي الإعاقات اللغوية. كما يمكن استخدام هذه الاختبارات مع

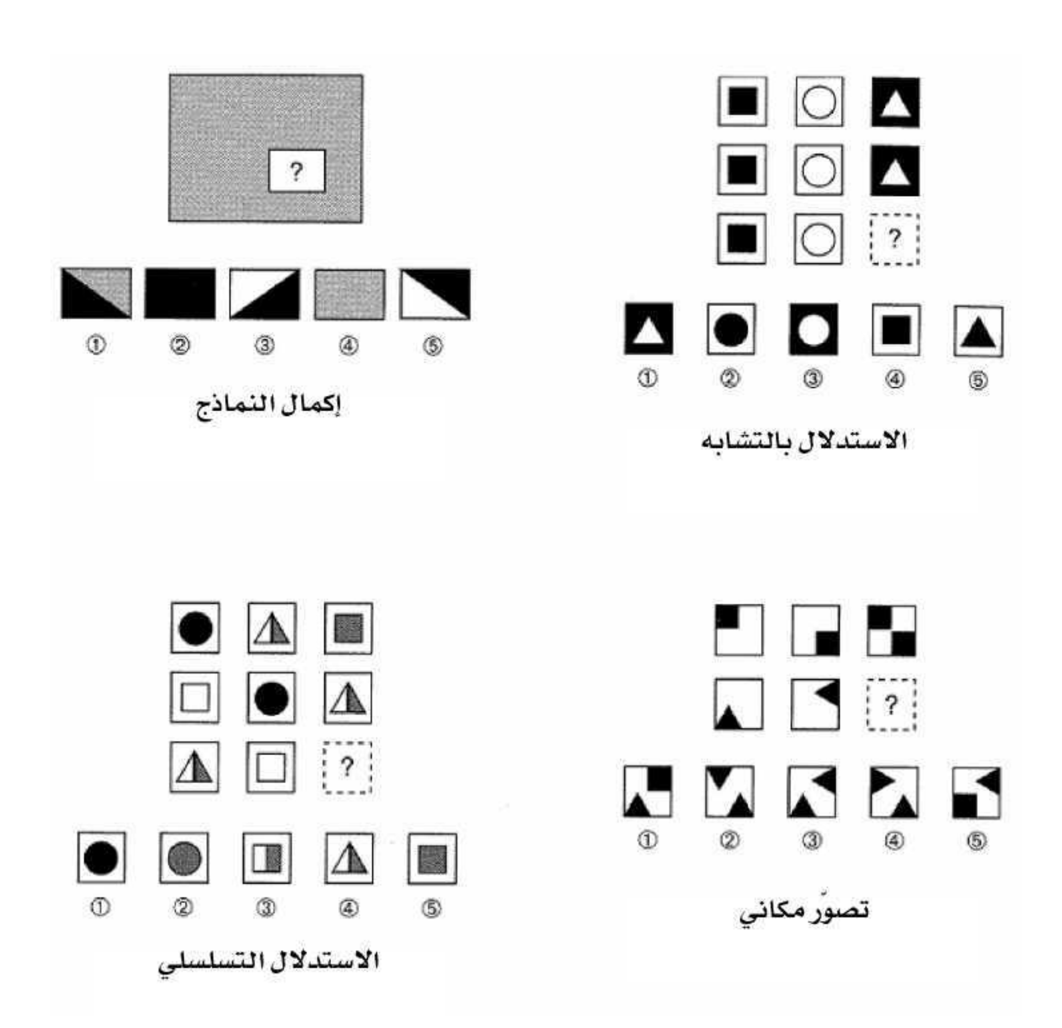
كافة الأفراد الذين لا ينتمون إلى الخلفيات الحضارية التي تسيطر عليها الحضارة الأمريكية واللغة الإنجليزية المستخدمة في بريطانيا وغيرها من بلاد الكومنولث.

لقد كان هدف مصممي الاختبارات، ولسنوات عديدة، هو وضع اختبارات ذكاء غير مرتبطة بحضارة معينة وذلك لاستخدامها في مجالات البحث والتطبيق العملي، إلا أن هذا الهدف سرعان ما تم تعديله ليركز على وضع اختبارات عادلة حضارياً لا يظهر فيها أي تحيز لفئة أو حضارة ما. ويعد اختبار مصفوفة ريفين التدرجية والذي تم تطويره في المملكة المتحدة واختبار الذكاء العادل حضارياً والذي تم تطويره في الولايات المتحدة من الأمثلة على الاختبارات العادلة حضارياً. يتألف اختبار الذكاء العادل حضارياً من ثلاثة مقاييس مقسمة على النحو التالي:

- المقياس ١: يستخدم هذا المقياس مع الأطفال بين أعمار ٤ و ٨ سنوات ومع البالغين المتخلفين عقلياً؛
- المقياس ٢: يستخدم هذا المقياس مع الأطفال بين أعمار ٨ و١٤ عاماً ومع البالغين ذوي مستويات الذكاء العادية؛
- المقياس ٣: يستخدم هذا المقياس مع طلاب الجامعات والمدراء والأشخاص الذين يعد ذكاؤهم فوق الطبيعي.

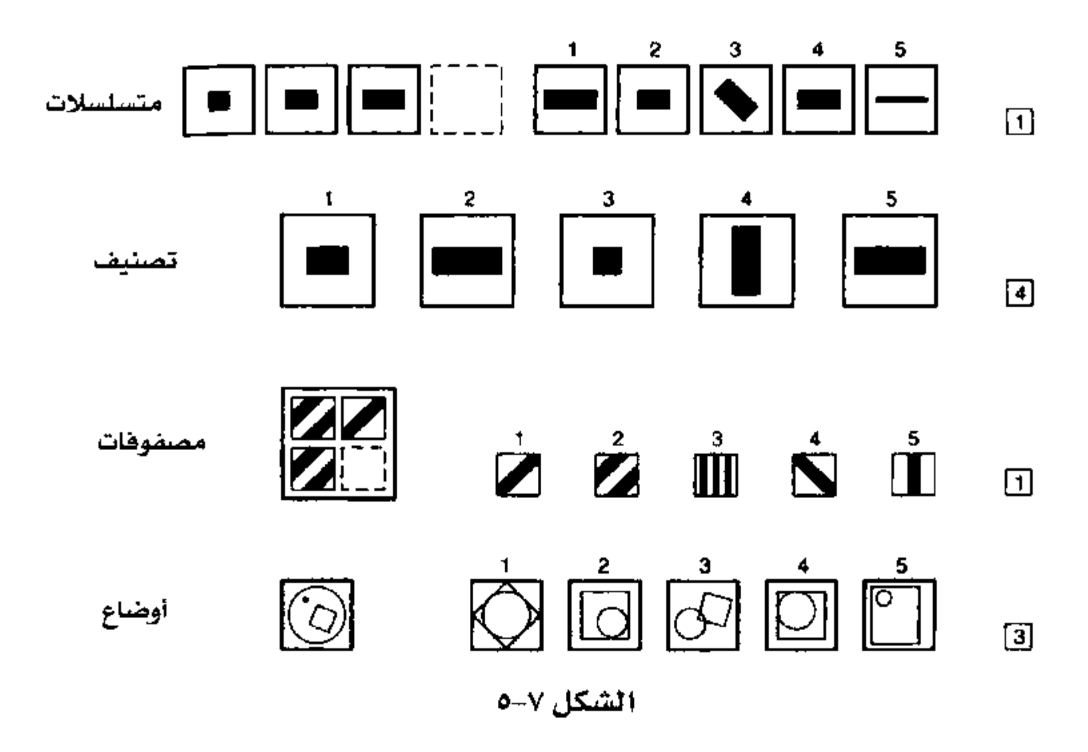
وكما يوضح الشكل ٧-٥، يتألف كل من مقاييس الاختبار من أربعة اختبارات فرعية (السلاسل والبنى والتصنيف والشروط)، ويضم المقياس ١، أربعة اختبارات فرعية إضافية تقيم المعلومات الحضارية عند الأفراد وإستيعابهم اللغوي. يستغرق الاختبار من ٤٠ إلى ٦٠ دقيقة للمقياس ١ ومن ٢٤ إلى ٣٠ دقيقة للمقياسين ٢ و٣.

على الرغم من الجهود التي بذلها مصممو الاختبارات ذوو النوايا الحسنة لإعداد اختبارات العادلة حضارياً، تبقى عملية وضع اختبارات ورقة وقلم على درجة مقبولة من الثبات والمصداقية تقيس القدرات الإدراكية للأفراد بمعزل عن التأثيرات الحضارية أمراً شبه مستحيل. فعلى سبيل المثال، يحصل البيض المنتمون إلى الطبقة الوسطى على درجات أعلى في اختبارات الذكاء اللغوية المليئة بالخصوصيات الحضارية من الأفراد الذين ينتمون لفئات عرقية وطبقات اجتماعية أخرى، كما أنهم يحصلون على درجات أعلى في الاختبارات العادلة حضارياً. إضافةً إلى ذلك، تعد الاختبارات اللغوية قادرة على التنبؤ بالتحصيل الأكاديمي والأداء الوظيفي وغيرها من مقومات النجاح في العالم الغربي بكفاءة أفضل من الاختبارات اللالغوية والاختبارات العادلة حضارياً.



الشكل ٧-٤

عينة من أسئلة اختبار ناغلييري اللالغوي للقدرة. (حقوق الطبع محفوظة لهاركورت بريس وشركاه. تمت إعادة الطبع بناء على موافقة الناشر. "NNAT" و"اختبار ناغلييري اللالغوي للقدرة" هي علامات مسجلة لمؤسسة علم النفس المتحدة).



الشكل ٧-٥

عينة من اختبار الذكاء العادل حضارياً. (حقوق الطبع ١٩٤٩، ١٩٦٠ لمركز اختبار الشخصية والقدرة. جميع الحقوق محفوظة. تمت إعادة الطبع مع الموافقة).

الخلاصة

على الرغم من الاهتمام الكبير الذي حظيت به الأبحاث التي قام بها غالتون في الأصول الوراثية للذكاء وطرق قياس وتقييم العبقرية، تم وضع أول اختبار للذكاء على يد بينيه وسيمون. يعد مقياس سيمون - بينيه للذكاء أداة قياس عمرية تم استخدامها بشكل أساسي في عملية التعرف على الأطفال المتخلفين عقلياً الذين يحتاجون إلى صفوف التعليم الخاص. وقد تم استخدام الاختبارات التي انحدرت من مقياس ستانفورد - بينيه في المجالات الأكاديمية ومجالات

الأعمال والجيش والدولة وغيرها من المؤسسات. كان اختبار ستانفورد - بينيه الأول للذكاء (والنسختين المعدلتين عنه) اختباراً عمرياً يحصل فيه الممتحن على عدد من الأشهر يمثل عدد الإجابات الصحيحة التي توصل إليها وتجمع هذه الأشهر لتعطي العمر العقلي للفرد الذي يتم تحويله بعد ذلك إلى معدل لذكاء هذا الفرد (IQ). وتختلف الطبعة الرابعة من اختبار ستانفورد - بينيه عن الطبعات الثلاثة الأولى من الاختبار وذلك في طريقة إجراء الاختبار وتصحيحه وفي كون هذه الطبعة تعتمد على نموذج هرمي للذكاء. تعتبر الاختبارات الأربعة في سلسلة ويكسلر WPPSI-R وWISC-III و WAIS-R و WAIS-III مقاييس نقطية (أو مقاييس درجات) يحصل الممتحَن فيها على عدد معين من النقاط لكل فقرة يجيب عليه بشكل صحيح ضمن اختبار فرعي ما. وتتألف جميع اختبارات ويكسلر من عدد من الاختبارات الفرعية التي تم تجميعها ضمن مقياسين أحدهما لغوي والآخر أدائي. يتم تحويل الدرجات الأولية في كل من الاختبارات الفرعية إلى درجات معيارية يتم جمعها لاحقاً للتعطينا معدل درجات معياري لغوي وآخر أدائي ومعدل درجات عام. وتحول هذه بدورها إلى معدلات ذكاء إنحرافية.

ومن اختبارات الذكاء الفردية الأخرى الجديرة بالذكر الاختبارات التالية:

- بطارية كوفمان لتقييم الأطفال (K-ABC)
- اختبار كوفمان القصير للذكاء (K-BIT)
- اختبار كوفمان لذكاء المراهقين والبالغين (KAIT)
- اختبارات سلوسون للذكاء (SIT)
- مقياس القدرة التمايزي (DAS)
- نظام داس - ناغلييري للتقييم الإدراكي (CAS)

وعلى الرغم مـن أن اختـباري CAS وK-ABC يسـتندان إلى نظرية لوريا - داس العصبيـة- النفسية للمعالجة العقلية المتزامنة والمتلاحقـة، يعود الفضل الأول في تصـميم العديد من الاختبـارات المذكورة أعلاه إلى البحث التجـريبي وليس إلى النظريات. وتسـتهدف بعض الاخـتبارات الأطفـال والمراهقين فـقط (K-ABC وDAS وCAS) بينمـا يسـتهـدف البـعض الآخـر المراهقين والبـالغين (KAIT) ويسـتهـدف البـعض الآخـر شـتى الفئـات العـمـرية من الطفـولة إلى الشيخوخة (K-BIT وSIT).

تتميز اختبارات الذكاء الجماعية بسهولة إجرائها وتصحيحها إلا أن هذه الاختبارات غالباً ما تكون أقل ثباتاً من الاختبارات الفردية. وتستخدم اختبارات الذكاء الجماعية كأدوات للاختيار الأولي بين المتقدمين في مجالات التوظيف والمقرارات والفصول الدراسية المختلفة. يعد اختبارا الجيش ألفا وبيتا من أول اختبارات الذكاء الجماعية استخداماً وقد تم تطوير عدد كبير من اختبارات الذكاء الجماعية التي تتألف من اختبارات فرعية منفصلة أو من أسئلة متعددة المستويات مرتبة ضمن كتيب لولبي. ولعل من أكثر هذه الاختبارات شعبية: اختبار القدرة الإدراكية واختبار أوتيس - لينون للقدرة المدرسية واختبار وندرليك للموظفين واختبار المهارات الإدراكية.

تم إعداد أدوات لقياس الذكاء تعتمد بشكل أساسي على الأداء وعلى فقرات لالغوية وذلك بسبب إنعدام الكفاءة اللغوية عند بعض الأشخاص في اللغة الإنجليزية التي تعد اللغة الأساسية في عدد كبير جداً من اختبارات الذكاء. ومن الاختبارات اللالغوية الفردية للذكاء اختبار غودينوف - هاريس للرسم واختبار الذكاء اللالغوي، بينما يمثل مقياس القدرة العامة للبالغين واختبار البنى المتشابهة واختبار ناغلييري للقدرة اللالغوية الاختبارات اللالغوية الجماعية للذكاء.

وعلى الرغم من فشل المحاولات في إعداد اختبارات غير متأثرة بالحضارة، تم تحقيق بعض النجاح في وضع اختبارات عادلة حضارياً كاختبار مصفوفات ريفين التدرجية واختبار الذكاء العادل حضارياً. ومع الأسف تنخفض نسبة الثبات في هذه الاختبارات عن نظائرها من الاختبارات المحملة بالتأثيرات الحضارية كما أن نسب الارتباط بين نتائج هذه الاختبارات ومعايير النجاح في الحضارة الغربية ضعيفة كذلك.

مصطلحات للمراجعة

راجع معاني المصطلحات المدرجة أدناه والتي تم استخدامها في هذا الفصل. الرجاء مراجعة فهرس التعريفات أو المعجم في حال وجود أي التباس حول معاني المصطلحات.

مقياس العمر	اختبار الذكاء الجماعي
العمر القاعدي	اختبار الذكاء الفردي
المستوى القاعدي	الذكاء
سقف الاختبار العمري	العمر العقلي
الاختبارات التكيفية	المعالجة العقلية للمعلومات
المحوسبة (CAT)	الاختبار متعدد المستويات
المستوى الحرج	اختبار لالفظي (لا لغوي)
القدرة المتبلورة	معدل الذكاء النسبي IQ
الاختبارات العادلة حضارياً	الوحدة الوظيفية الثانية
معدل الذكاء IQ الإنحرافي	المعالجة التسلسلية (المتابعة) للمعلومات
علم طبيعة المعرفة	المعالجة المتزامنة للمعلومات
الوحدة الوظيفية الثانية	الوحدة الوظيفية الثالثة
القدرة الفطرية	اختبار لغوي (لفظي)

المراجع المقترحة

Daniel, M. H. (1997). Intelligence testing: Status and trends. *American Psychologist, 52,* 1038–1045.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269,* 41–45.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve.* New York: Free Press.

Matarazzo, J. D. (1992). Psychological testing and assessment in the 21st century. *American Psychologist, 47,* 1017–1018.

Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science, 1,* 86–59.

Schaie, K. W. (1994). The course of adult intellectual development. *American Psychologist, 49,* 304–313.

Shurkin, J. N. (1992). *Terman's kids.* Boston: Little, Brown.

Suzuki, L. A., & Valencia, R. R. (1997). Race-ethnicity and measured intelligence: Educational implications. *American Psychologist, 52,* 1103–1114.

Vernon, P. A. (Ed.). (1993). *Biological approaches to the study of human intelligence.* Norwood, NJ: ABLEX.



تقييم التطور والاضطرابات الإدراكية

تقوم اختبارات الذكاء الفردية والجماعية بعمل تقييم القدرات الإدراكية عند الأفراد المنتمين إلى فئات عمرية مختلفة بشكل جيد، وتختلف هذه الاختبارات من حيث مستويات سقف وقاعدة القدرات الإدراكية التي تختبرها. فعلى حين تستعمل بعض الاختبارات مع المستويات العمرية والإدراكية المتدنية، ينحصر عمل بعض الاختبارات الأخرى مع الأفراد الأكبر عمراً والأقدر إدراكاً إذ لا يسمح حد العمر الأدنى لاختبار ستانفورد - بينيه ولسلسلة ويكسلر باختبار الأطفال دون سن الثالثة أو الأفراد الذين يعانون من التخلف العقلي الشديد. وفي الوقت ذاته لا يسمح الحد الأعلى لهذين الاختبارين بتقييم الأفراد الموهوبين. يتطلب تقييم هذه النوعية الخاصة من الأفراد وإعداد دراسات تشخيصية للتطور اختبارات وإجراءات خاصة تعنى بالنواحي الجسدية والإدراكية واللغوية والعاطفية وبالتطور الذاتي للأفراد.

وقد تم تصميم اختبارات لالغوية واختبارات أدائية واختبارات عادلة حضارياً، تستخدم أساساً مع الأطفال والبالغين من ذوي الإعاقات اللغوية

والتعليمية (وهو ما تحدثنا عنه سابقاً في الفصل السابع من الكتاب)، من المطلوب كذلك توفر وسائل لتقييم الأفراد ذوي الإعاقات الجسدية، إذ لا يمكن تقييم القدرات الإدراكية للأطفال الذين يعانون من إعاقات بصرية أو سمعية أو حركية باستخدام اختبارات الذكاء التقليدية. ومن الممكن أن تؤدي الإعاقات الحسية والحركية إلى صعوبات أو تأخر في التعلم عند بعض الأطفال كما يمكن أن تنتج صعوبات التعلم عن حالات عصبية منشؤها تغيرات واضحة أو مخفية في الجهاز العصبي. ولهذا السبب يتم استخدام اختبارات خاصة للذاكرة والإدراك الحسي والمهارات الحركية - النفسية وغيرها من القدرات إلى جانب المقاييس العامة للقدرة العقلية وذلك لتكوين صورة تشخيصية أكثر وضوحاً عن الأطفال المتأخرين داخل وخارج المدرسة.

لا تستخدم الاختبارات التي يتناولها هذا الفصل بالشكل الواسع الذي تستخدم به اختبارات الذكاء التقليدية، إلا أن هذه الاختبارات تعتبر أدوات وإجراءات إضافية لفهم الأطفال والبالغين الذين يواجهون صعوبات في التأقلم مع متطلبات الحياة اليومية ولتصميم برامج تعليمية مناسبة وإتخاذ خطوات علاجية تخدم مصالحهم.

اختبارات الارتقاء عند الرضع والأطفال الصغار

تعود جذور الاهتمام بتطور الأطفال العقلي إلى القرن التاسع عشر الذي شهد اهتماماً ملحوظاً بمصلحة الأطفال قاد إلى سن قوانين وتشريعات في مجالات صحة وتعليم وعمالة الأطفال بالإضافة إلى المنهج العام للتعامل معهم. وقد تم نتيجة "لحركة رعاية الطفل" هذه تصميم عدد من أدوات القياس التي تقيم القدرات الإدراكية والحركية والإدراكية الحسية والعاطفية وقدرة النمو الاجتماعي عند الرضع والأطفال الصغار وذلك لضمان توفر إجراءات موضوعية لتشخيص الحالات الفردية وتدعيم البحث العلمي في مجال الإرتقاء الفكري عند الأطفال.

جداول جيزل الارتقائية

بدأ أرنولد جيزل من جامعة ييل في العشرينيات من القرن الماضي عمليات بحث منتظمة في النمو الإرتقائي عند الرضع والأطفال الصغار، وقد استندت أبحاث جيزل إلى مبادئ نظرية مفادها أن الإرتقاء الحركي (الذي يشمل الحركات الدقيقة وغير الدقيقة) واللغوي والذاتي - الاجتماعي والسلوك التكيفي عند الأفراد يتم بطريقة متسلسلة نضجياً (أي متتابعة وفقاً لعمر الفرد). وقد جمع جيزل والعاملون معه كماً هائلاً من المعلومات عن تطور عينة كبيرة من الأطفال وذلك عن طريق السجلات المنزلية للأطفال وتاريخهم المرضي وسجلاتهم اليومية وقياساتهم التشريحية Anthropometry وملاحظات الأمهات وتقارير العيادات النفسية عن سلوك الأطفال والاختبارات المعيارية وتقديرات الإرتقاء الفكري. وقد قادت هذه الأبحاث إلى وضع جداول جيزل الإرتقائية التي يحصل فيها الطفل على درجات بناءً على وجود أو غياب نمط معين من السلوك يعتبر طبيعياً عند الأطفال من نفس عمره. يتم تحويل هذه الدرجات إلى العمر الإرتقائي developmental age للطفل (DA) والذي يقودنا إلى المعدل الإرتقائي للطفل developmental quotient (DQ) عند تقسيمه على العمر الفعلي لهذا الطفل.

ويعتبر نشر الطبعة الثالثة من جداول جيزل بعد إنقضاء نصف قرن على ظهور طبعتها الأولى دليلاً على شعبية هذا الاختبار (كنوبلوخ وباسامانيك، ١٩٧٤؛ كنوبلوخ وستيفينز ومالون، ١٩٨٧). ولا تزال النسخ المعدلة عن الجداول الأصلية مستخدمة حتى يومنا هذا وبخاصة من قبل أطباء الأطفال[١]. أما أخصائيو علم النفس الارتقائي developmental psychologists فقد وضعوا أدوات قياس خاصة بهم تتصف بخصائص سايكومترية أفضل من تلك التي تملكها جداول جيزل.

---

(١) تعد مقاييس سمات تطور الأطفال (المتوفرة لدى شركة أنظمة علم السلوك، ص.ب. ٥٨٠٢٧٤، مينيابوليس، MN 55458) مثالاً على الاختبارات التي تم تعديلها عن جداول جيزل للتطور.

دفعت أبحاث جيزل وزملائه العديد من الباحثين إلى نشر اختباراتهم الخاصة بقياس ذكاء الرضع ومن الأمثلة على هذه الاختبارات: مقياس كاليفورنيا العقلي للسنة الأولى من عمر الطفل واختبار نورثوسترن للذكاء ومقياس غريفيث للتطور العقلي ومقياس كاتل لذكاء الرضع (وهو الاختبار الوحيد من هذه الفئة الذي لا يزال يطبع حالياً).

دنفر ٢

تم تصميم اختبار دنفر ٢ Denver-II، والذي يعد نسخة مراجعة عن جداول جيزل الإرتقائية، بغرض التعرف على تأخرات التطور عند الأطفال من عمر الولادة وإلى عمر ٦ سنوات. ويتألف الاختبار من ١٢٥ فقرة تستغرق من ٢٠ إلى ٢٥ دقيقة (ومن ١٠ إلى ١٥ دقيقة للنسخة المختصرة عن الاختبار). تقسم درجات الاختبار على أربع نواحي أساسية هي: النمو (الإرتقاء) الذاتي – الاجتماعي والحركي الدقيق –التكيفي واللغوي والحركي العام. كما يمكن الحصول على خمسة تقديرات للسلوك من الاختبار هي: الاعتيادي والمنفذ للتعليمات والمهتم بما حوله والذي لا يخشى شيئاً ومدى الانتباه. يتميز اختبار Denver-II بسهولة إجرائه ووضع درجاته ومما لا شك فيه أن هذا الاختبار أفضل بكثير من سابقه إلا أنه يعاني من بعض المشاكل السايكومترية ولا سيما تلك المتعلقة بكون عينة التقنين غير ممثلة للمجتمع (هيوز، ١٩٩٥).

مقياس بايلي لنمو الرضع

يعد مقياس بايلي لتطور الأطفال واحداً من أكثر اختبارات تقييم الذكاء عند الرضع والأطفال الصغار استخداماً اليوم (انظر الشكل ٨–١). وقد تم تصميم الطبعة الثانية من هذا الاختبار (BSID-II) والتي نشرت في عام ١٩٩٣ لتقييم

الأطفال بين أعمار شهر و٣٦ شهراً والذين هنالك شك في أنهم "في خطر[٢]". يستغرق الاختبار من ٢٥ إلى ٣٥ دقيقة مع الأطفال دون سن ١٥ شهراً وإلى مدة أقصاها ٦٠ دقيقة مع الأطفال فوق هذا السن.



الشكل ٨-١

صورة لمقياس بايلي لنمو الرضع - الطبعة الثانية: (تم النقل بإذن من المؤسسة النفسية ١٩٩٣، إن مقياس بايلي لنمو الرضع هو علامة مسجلة للمؤسسة النفسية).

(٢) لا يوجد تعريف دقيق لمصطلح "في خطر" "at risk" إلا أنه يشير بشكل عام إلى الأطفال الذين يعانون من تأخر في النمو الذهني أو من إعاقة جسدية تحتاج إلى إجراءات علاجية. وفي الواقع العملي يعتبر الأطفال دون سن المدرسة "في خطر" إذا كانوا غير مناسبين للإلتحاق بالصف الأول أو إذا كانوا يعانون من صعوبات في التعلم لا تتضح من خلال الملاحظة اليومية بل يحتاج الكشف عنها إلى عملية مسح نفسي أو اختبار فردي.

يتألف اختبار BSID-II من ثلاثة أقسام هي: مقياس عقلي يعطي مؤشرعن التطور العقلي للطفل ومقياس حركي يعطي مؤشراً عن التطور الحركي – النفسي له وسلم لتقدير السلوك يوفر معلومات تكميلية للمقياسين الحركي والعقلي. يقوم المقياس العقلي بقياس ما يلي:

- حدة الإحساس والإدراك الحسي
- التمييز
- القدرة على الإستجابة إلى ما يميزه العقل
- اكتساب الطفل للقدرة على تثبيت الأشياء في ذاكرته
- الذاكرة
- التعلم
- حل المشاكل
- النطق
- بدايات التواصل اللغوي
- الدلائل الأولى على عملية التفكير التجريدي
- التعود
- التخطيط الذهني (Mental mapping)
- اللغة المعقدة
- تكوين المفاهيم الرياضية.

أما المقياس الحركي فيقيس:

- القدرة على التحكم بالجسم
- التناسق بين العضلات الكبيرة
- الحركات الدقيقة للأصابع واليدين
- الحركة الديناميكية
- التطبيق الديناميكي
- المحاكاة الحركية
- التحليل اللمسي للأشياء.

وأخيراً، يقيس سلم تقدير السلوك الانتباه/الاثارة والتوجه/الإستغراق والضبط العاطفي وجودة الحركة.

تم قياس اختبار BSID-II عام ١٩٩٠ على عينة عشوائية مؤلفة من ١٧٠٠ طفلاً تم تقسيمها طبقياً تبعاً للمنطقة الجغرافية والعمر والجنس والعرق ودرجة تعليم الوالدين. ويضم دليل المستخدم للاختبار معلومات عن العينات السريرية للأطفال. ويرافق اختبار BSID-II مسّاح بايلي للإرتقاء العصبي للرضع (BINS) Baley Infant Neurodevelopmental Screen المصمم لتقييم الوظائف العصبية والسمعية والبصرية الإستقبالية والعمليات الاجتماعية الإدراكية عند الأطفال الذين تتراوح أعمارهم بين ٣ و ٢٤ شهراً. ويبدو أن قياس هذا الاختبار قد تم كذلك بشكل جيد إلا أن عينة القياس تبقى صغيرة بعض الشيء.

خصائص اختبارات ذكاء الرضع

من الطبيعي جداً أن اختبار الأطفال دون عمر المدرسة هو أمر شديد التعقيد مقارنة بالأطفال في العمر المدرسي وذلك بسبب قصر مدى الانتباه عند الأطفال في هذه المرحلة وتعرضهم للتعب في فترة قصيرة وانخفاض الدافع

لديهم على إتمام المهام الموكلة إليهم. ونتيجة لذلك تنخفض نسب ثبات وصدق الاختبارات المصممة لهذه الفئة العمرية كما تنخفض نسب الارتباط بين درجات الأطفال في هذه الاختبارات وأدائهم المدرسي أو نتائجهم في اختبارات ذكاء خضعوا لهافي سنواتٍ لاحقة. كما تختلف الاختبارات المصممة للأطفال دون سن الثالثة عن تلك التي تستهدف الأطفال من عمر المدرسة من حيث نوعية المهام المطلوبة. وعلى حين تتألف اختبارات ستانفورد - بينيه واختبار WISC-III من فقرات لغوية ومهام مشابهة للمهام المدرسية، تتألف اختبارات ذكاء الرضع حصراً من فقرات حسية - حركية. هذا يعني أن اختبارات ذكاء الرضع قد تساعد في التعرف على حالات التخلف العقلي والتلف الدماغي الولادي إلا أنها لا تميز بين القدرات الإدراكية للأطفال التي تحدد النجاح المدرسي والعملي.

اختبارات أخرى للأطفال الصغار

يعتبر مقياس ويكسلر لذكاء الأطفال دون سن المدرسةوأطفال المرحلة الإبتدائية (الذي يُعنى بالأطفال من ٣ إلى ٧ والذي تحدثنا عنه في الفصل السابع من الكتاب) (WPPSI-R) ومقياس ماكارثي لقدرات الأطفال (MSCA) من أهم الاختبارات الفردية المصممة للأطفال الذين لم يدخلوا المدرسة. يتابع MSCA عمل اختبار BSID-II إذ يختبر الأطفال الذين تتراوح أعمارهم بين ٢,٦ إلى ٨,٦ عاماً، وعلى الرغم من قدمه (معاييره تعود إلى عام ١٩٧٢) يبقى اختبار MSCA مفيداً في تحديد القدرات الذهنية العامة عند الأطفال. يستغرق الاختبار بين ٤٥ إلى ٦٠ دقيقة ويمكن وضع درجاته بالنسبة إلى ستة مقاييس هي: المقياس اللغوي، والمقياس الإدراكي الحسي - الأدائي، والمقياس الكمي، والمقياس الإدراكي العام (وهو مجموع المقاييس الثلاثة السابقة)، ومقياس الذاكرة، ومقياس الحركة.

وقد نص قانون الأفراد ذوي الإعاقات التعليمية (IEDA) على ضرورة خضوع الاختبارات التي تقيم أعداد كبيرة من الأطفال من حيث التأخر في النمو والاختبارات التشخيصية التفصيلية إلى عدد من المعايير. وعلى الرغم من أن اختبار MSCA يطابق هذه المعايير إلا أن الاختبارالمسّاح لتقييم مظاهر حياة الأطفال دون سن المدرسة FirstSTEP واختبار خدمات التوجيه الأمريكية AGS لمظاهر المسح المبكر، هما الاختباران اللذان تم إعدادهما بناءً على توصيات IEDA. يعد اختبار FirstSTEP اختبار مسح سريع يستغرق ١٥ دقيقية ويكشف عن تأخر النمو العقلي عند الأطفال بين ٢,٩ إلى ٦,٢ عاماً. يتألف الاختبار من ١٢ اختباراً فرعياً تهدف إلى خلق جو مرح ولعب في الاختبار، وقد تم تصنيف هذه الاختبارات الفرعية إلى ثلاثة من النواحي الخمسة التي نصت عليها توصيات IEDA وهي: الإدراك، والاتصال، والحركة. يتم حساب درجات الفرد في الاختبارات الفرعية الإثنتي عشرعلى شكل درجات مركبة يتم تصنيفها إلى ثلاث فئات هي: "ضمن الحدود المقبولة" و"توخي الحذر" (أي هنالك دليل على تأخر بسيط إلى متوسط في النمو العقلي) و"في خطر" (معرض للتأخر في النمو العقلي). وتتوفر سلالم تقدير اختيارية اجتماعية – عاطفية للمدرسين/الوالدين تستخدم لتقييم الناحية الرابعة التي نص عليها IEDA (وهي: مستويات الانتباه/الحركة، والتفاعل الاجتماعي، والخصائص الشخصية، ومشاكل السلوك الكبيرة)، بينما تساعد سلالم تقدير اختيارية كذلك للسلوك التكيفي في تقييم الناحية الخامسة من IEDA (وهي: نشاطات الحياة اليومية والسيطرة على النفس والعلاقات والتفاعلات الاجتماعية والتصرف داخل المجتمع).

ومن الاختبارات السريعة التي تساعد على تحديد التأخر في الإرتقاء العقلي عند الأطفال دون سن المدرسة (أي الذين تتراوح أعمارهم بين ٢ إلى ٦ أعوام و٧ أشهر) اختبار AGS لمظاهر المسح (ESP) والتي تتألف من ثلاثة مظاهر أساسية تستغرق أقل من ٣٠ دقيقة وأربعة استطلاعات (إضافية أو

تكميلية) تستغرق بين ١٥ إلى ٢٠ دقيقة. ويتألف مظهر الإدراك/اللغة من عدد من المهام التي تقيم مهارات الاستدلال والترتيب البصري والتمييز ولغة الإستقبال والتعبير ومهارات المدرسة الأساسية. ويقيم مظهر الحركة القدرة الحركية العامة والدقيقة (كالسير ضمن خط مستقيم وتقليد حركات الذراعين والرجلين وتتبع المتاهات بالقلم ورسم الأشكال). أما مظهر مساعدة الذات أو المظهر الاجتماعي فيتألف من استبيان يملؤه أحد الأبوين أو المسؤول الرئيس عن رعاية الطفل ويهدف إلى جمع معلومات حول أداء الطفل الإعتيادي في الإتصال مع الآخرين ومهارات الحياة اليومية والإختلاط مع الناس والمهارات الحركية. أما الاستطلاعات الملحقة باختبار ESP فهي: استطلاع النطق (الذي يتطلب من الطفل أن ينطق بعشرين كلمة)، واستطلاع البيئة المنزلية (والذي يجيب فيه الأب أو الأم عن أسئلة حول البيئة المنزلية للطفل)، واستطلاع التاريخ الصحي (والذي يضع فيه الوصي على الطفل إشارة أمام المشاكل الصحية التي يعاني منها الطفل)، واستطلاع السلوك (والذي يقيم فيه القائم على الاختبار سلوك الطفل أثناء إجراء مظهر الإدراك/اللغة ومظهر الحركة من حيث مدى انتباهه وقدرته على تحمل الإحباط عند الفشل في الإجابة وأسلوبه في الإجابة). يتم تحويل درجات اختبار ESP إلى مؤشرات أولية (على المستوى الأول) للمسح أو إلى درجات قياسية ورتب مئوية ومكافئات عمرية على المستوى الثاني تحدد ما إذا كان هذا الطفل يحتاج إلى تقديرات أخرى. وعلى الرغم من أن الخصائص السايكومترية لاختباري ESP وFirstSTEP تعتبر مقبولة بالنسبة لأدوات المسح الإرتقائي، لم يتم استخدام هاتين الأداتين بشكل واسع في عمليات البحث النفسي.

اختبار الأفراد المعاقين جسدياً:

يعد الأفراد الذين يعانون من التخلف العقلي أكثر عرضة من غيرهم للصعوبات الحركية والإضطرابات الجسدية، إلا أن هذا لا يعني أن الصعوبات

الحركية تنحصر بالمتخلفين عقلياً. ونظراً لتأثير هذه الصعوبات على أداء الأفراد في اختبارات القدرة الإدراكية، وجد رواد الاختبارات النفسية أنفسهم وجهاً لوجه مع التأثيرات المحيِّرة للصعوبات الجسدية والعقلية على السلوك البشري.

ومن الضروري جداً التمييز بين الصعوبة (Handicap) أو الإعاقة (Disability) والعجز (Impairment)[٣] إذ تؤدي الصعوبات في الحركة إلى درجات متفاوتة من الإعاقة والعجز تختلف ليس باختلاف طبيعة الصعوبة فحسب بل باختلاف شخصية الفرد المتأثر بهذه الصعوبة. فمما لا شك فيه أن صعوبة الحركة تجعل مهمة الفرد في التعامل مع الحياة شاقةً ومعقدة إلا أنها بدلاً من أن تقود إلى عجز دائم في الوظائف السلوكية عند هذا الفرد يمكن أن تتحول إلى تحدٍ يجب اجتيازه. ومن الجائز جداً أن التعويض عن الصعوبات في التعلم والحركة التي يعاني منها شخص ما قد يقود هذا الشخص إلى الإرتقاء بأدائه إلى درجات لم يكن يخطر في باله أن يصل إليها. وهذا يعني أنه من المهم جداً عدم تركيز كافة الاهتمام على التعرف على صعوبات التعلم والحركة التي يعاني منها فرد ما بل على سلوكه التكيفي في التعامل مع نقاط ضعفه. وتتوفر قوائم للرصد وسلالم للتقدير تقييم السلوك التكيفي عند الأفراد ويمكن الحصول على معلومات عنها من الجزء الأول من هذه السلسلة (أيكين، ١٩٩٦، ص١٨٢-١٧٨). وتضم السلوكات التي تم تقديرها بأدوات التقييم هذه مجموعة من النشاطات والمهام المرتبطة بالحياة اليومية التي يقوم بها الأفراد لكي يتأقلموا مع البيئة المحيطة بهم.

---

(٣) على الرغم من أن كلمة Handicap تترجم عادةً بكلمة "إعاقة" من الواضح جداً هنا أن الكاتب يريد التمييز بين عدم القدرة على فعل شيء ما (Disability) والتعرض إلى صعوبات عند القيام بعمليات عقلية أو جسدية معينة. ونظراً للمعنى التاريخي لكلمة Handicap والذي يعني فرض صعوبات على فريق ما في الألعاب لجعل فرص الربح للفرق الأخرى أكثر عدالة، رأيت أنه من الأنسب ترجمة كلمة handicap على أنها صعوبة وليست إعاقة. (المترجم).

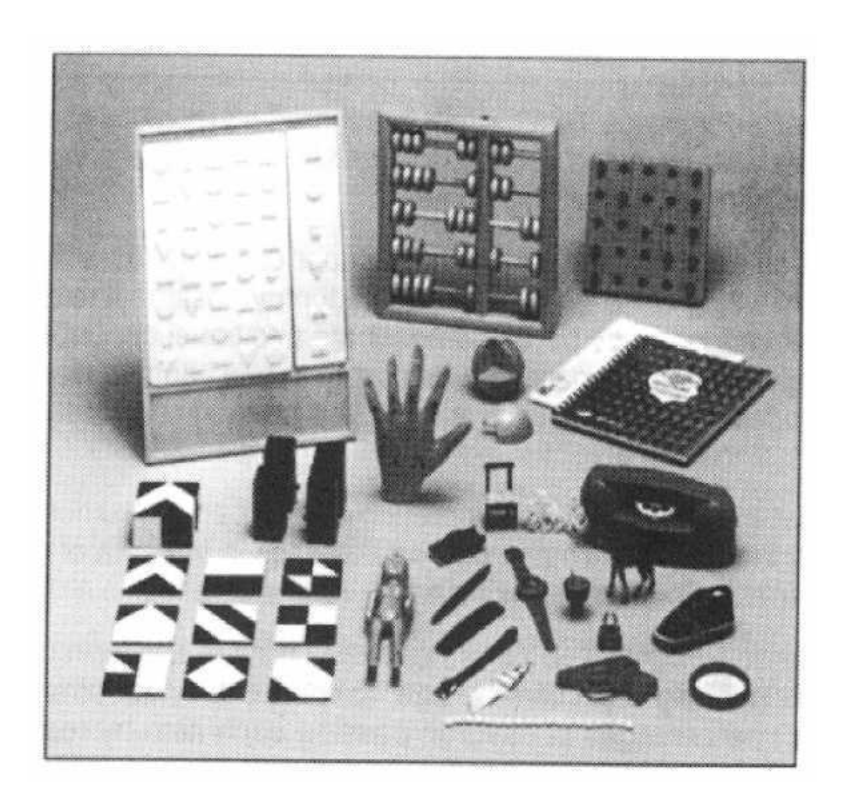
وتؤدي صعوبات الحركة والسمع والبصر وغيرها إلى عدم وصول الذين يعانون من هذه الصعوبات إلى المستوى الذي يتناسب مع إمكاناتهم العقلية في اختبارات الذكاء التقليدية مثل اختبار ستانفورد - بينيه وسلسلة ويكسلر. وقد قاد هذا إلى وضع عدد من الاختبارات الخاصة التي تضع بعين الاعتبار الصعوبات الحركية والحسية للممتحَنين، ويمكن الحصول على قائمة تمثيلية لهذه الاختبارات باختيار الفقرة ٤ من قائمة برنامج H-7 المتوفر على القرص المرافق لهذا الكتاب.

اختبارات المكفوفين وضعيفي البصر

تم تعديل عدد من اختبارات الذكاء التقليدية كما تم وضع اختبارات خاصة وذلك لتقييم القدرات الإدراكية للذين يعانون من ضعف البصر أو فقده. فقد ظهر اختبار هايز - بينيه وهو نسخة معدلة عن اختبار ستانفورد - بينيه أعدت للأطفال المكفوفين وضعيفي البصر بعد ظهور الاختبار الأصلي بوقت قصير، كما ظهر اختبار آخر في وقت لاحق هو اختبارات بيركينز - بينيه لذكاء المكفوفين (دايفيز، ١٩٨٠). وعلى حين تم إيقاف نشر هذين الاختبارين لا يزال المقياس اللمسي (Haptic) لذكاء البالغين فاقدي البصر واختبار الاستعداد التعليمي للمكفوفين متوفرين في الأسواق.

تشير كلمة Haptic إلى المعلومات التي يتم الحصول عليها عن طريق اللمس وهي الطريقة التي، بالإضافة إلى حاسة السمع، يتعلم عن طريقها المكفوفون ويتم اختبارهم. إن المقياس اللمسي لذكاء البالغين فاقدي البصر هو أداة لقياس ذكاء الأفراد الذين يزيد عمرهم عن ١٦ عاماً ويمكن استخدام هذا الاختبار بمفرده أو مع مقياس ويكسلر لذكاء البالغين (انظر الشكل ٨-٢). يعتمد هذا الاختبار على محفزات لمسية بدلاً من المحفزات البصرية التي تستخدمها اختبارات الذكاء التقليدية ويطلب من ضعاف البصر ارتداء عصبة لتغطية

عينيهم. وتستغرق الاختبارات الفرعية التي يتألف منها الاختبار (وهي: الرموز الرقمية، وتجميع الأشياء، وتصميم المكعبات، وإتمام الأشياء، ولوح الأنماط، والحساب عن طريق محسب الخرز) بين ٦٠ إلى ٩٠ دقيقة.



الشكل ٨-٢

المواد المستعملة في المقياس اللمسي للذكاء (تمت إعادة النشر بموافقة شركة ستولتينغ)

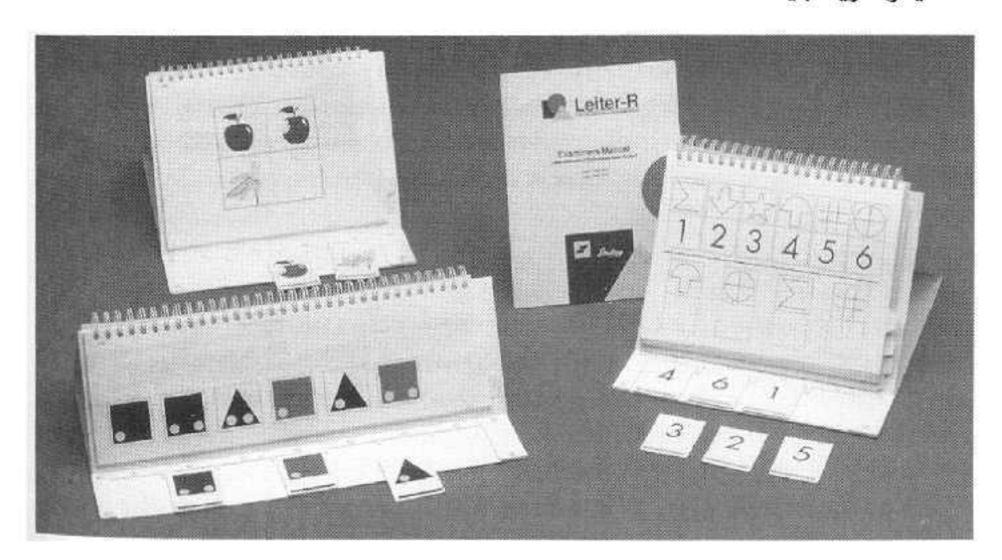
يشابه اختبار القابلية التعليمية للمكفوفين (BLAT) المقياس اللمسي لذكاء البالغين من حيث اعتماده على مهام لمسية إلا أنه يختلف عنه في أنه يستهدف الأطفال والمراهقين بين أعمار ٦ و٢٠ عاماً. وتتألف مهام هذا الاختبار غير المحدود زمنياً من تعديلات لمسية لمهام مصفوفات ريفين التدرجية ومهام اختبار

الذكاء العادل حضارياً إذ تم تحويل جميع الفقرات الاختبارية إلى نقط بارزة وخطوط محفورة تشابه كتابة بريل. ويعطي أداء الفرد في الأنواع الستة لمهام الاختبار معدل تعلمه الذي يساوي وسطه الحسابي ١٠٠ وإنحرافه المعياري ١٥. تم قياس اختبار BLAT على عينة طبقية صغيرة من الطلاب المكفوفين وعلى الرغم من أن معاملات الإتساق الداخلي وثبات إعادة الاختبار مقبولة، لا تتوفر معلومات في دليل المستخدم لاختبار BLAT عن نسب ارتباط الاختبار مع غيره من الاختبارات باستثناء نسبة ارتباط ٠,٧٤ مع اختبار هايز – بينيه ونسبة ارتباط ٠,٧١ ومع اختبار WISC اللغوي.

اختبار الصم وضعيفي السمع

يفضل العاملون في مجال الاختبارات النفسية مقياس الأداء من سلسلة ويكسلر عند اختبار الأفراد الذين يعانون من صعوبات شديدة في السمع، والمقياس اللغوي من نفس السلسلة للمكفوفين وضعيفي البصر. ويتوفر عدد من التعديلات على اختبارات سلسلة ويكسلر لجعلها أكثر مناسبة للصم وضعيفي السمع كالنسخة المعدلة للاستخدام مع الأطفال الصم عن اختبار ويكسلر للأداء WISC-R. ومن الاختبارات المناسبة للاستخدام مع ضعاف السمع، الاختبارات اللالغوية كاختبار Leiter-R وهو مقياس عادل حضارياً لتقييم القدرة الإدراكية لدى الأفراد من شتى الخلفيات الحضارية. يستهدف الاختبار الأفراد بين ٢ إلى ٢١ عاماً ويمكن، نظراً لطبيعته اللالغوية، استخدامه ليس مع الأطفال الصم فحسب بل مع جميع الأطفال الذين يعانون من مشاكل لغوية (الأطفال غير الناطقين باللغة الإنجليزية والذين يعانون من اضطراب نقص الانتباه وفرط الحركة ADHD والذين يعانون من الإعاقات التعليمية والذين تعرضوا إلى إصابات دماغية مؤلمة والذين ينتمون إلى حضارات أخرى والمعاقون حركياً والذين يعانون من مرض التوحد) أو مع الأطفال الموهوبين. ويتطلب الاختبار من

الممتحَنين المطابقة بين عدد من بطاقات الإجابة وعدد من الرسومات الموجودة على لوح يشبه ذلك الذي يستخدمه الرسامون (انظر الشكل ٨-٣). وتستغرق اختبارات الإستدلال الفرعيةالأربعة واختبارات التصور الفرعية الستة التي تؤلف بطارية الإستدلال والتصور ٤٠ دقيقة على حين تستغرق اختبارات الذاكرة الفرعية الثمانية واختبارا الانتباه الفرعيين الذين يؤلفون بطارية الذاكرة والانتباه ٣٥ دقيقة فقط.ويمكن إجراء مسح سريع لمعدل الذكاء أو إعاقات التعلم أو اضطراب نقص الانتباه وفرط الحركة ADHD في ٢٥ دقيقة وذلك باستخدام أجزاء من البطاريات. ويستغرق تقييم الموهوبين ٣٥ دقيقة. تم قياس اختبار Leiter-R في عام ١٩٩٣ على عينة مؤلفة من ١٧١٩ فرداً طبيعياً و٦٩٢ فرد غير عادي تتراوح اعمارهم بين ٢ و٢١ عاماً. وتشير المعلومات التي يحتويها دليل المستخدم عن ثبات وصدق الاختبار إلى كون اختبار Leiter-R أداة قياس سايكومترية جيدة.



الشكل ٨-٣

عينة من المواد المستخدمة في بطاريات التصور والاستدلال والانتباه والذاكرة من اختبار Leiter-R (تمت إعادة الطبع بموافقة شركة ستولتينغ)

تم تصميم اختبار هيسكي - نبراسكا لقابلية التعلم حصراً للأطفال (بين أعمار ٣ و١٧) الذين يعانون من اضطرابات في السمع، ويتألف هذا الاختبار من الاختبارات الفرعية اللالغوية التالية:

أنماط الخرز	طوي الأوراق	تذكر الأرقام
تذكر الألوان	مدى الانتباه البصري	المكعبات الناقصة (Puzzle)
التعرف على الصور	أنماط المكعبات	الصور المتشابهة
الربط بين الصور	إكمال الرسوم	الإستدلال المكاني

يمكن تقديم اختبار هيسكي - نبراسكا للأطفال الصم عن طريق الإيماء وللأطفال ذوي السمع الطبيعي عن طريق القراءة، ويتم التعبير عن أداء الأطفال في الاختبار عن طريق عمر التعلم LA ومعدل التعلم LQ وذلك عند استخدام الإيماء، أو عمر عقلي MA ومعدل ذكاء IQ عند تقديم الاختبار بالشكل العادي. ومع الأسف، تعود معايير الاختبار إلى أكثر من ٣٠ عاماً خلت ولا يحتوي دليل المستخدم للاختبار الذي صدر عام ١٩٦٦ على أية معلومات عن الخصائص السايكومترية للنسخة المراجعة من هذا الاختبار.

إن صعوبات اختبار الأفراد الصم او المكفوفين تتضاعف عند اجتماع الإعاقتين في الفرد ذاته، وفي هذه الحالة يتم استخدام سلم لتقدير السلوك أو قائمة رصد له بدلاً عن الاختبارات القياسية. ومن الأمثلة على هذا النوع من أدوات القياس مقياس كاليه - آزوسا وهو قائمة رصد لتقييم سلوك الأطفال (بين أعمار ٠ إلى ٩ سنوات) الصم أو المكفوفين على مقياسين فرعيين في عدد من النواحي (الإدراكية الحسية، واللغة، والحياة اليومية، والمهارات الاجتماعية...إلخ). ولا يتم وضع إشارة أمام سلوك معين للطفل إلا في حال كون هذا السلوك "موجوداً تماماً ومنتظماً". ويمكن للوالدين أو المدرسين أو غيرهم

ملء مقياس كاليه – آسوزا بشكل منتظم وذلك لتقييم تطور الأطفال الخاضعين لبرنامج علاجي ما.

صعوبات الحركة ذات المنشأ العظمي

نظراً للبطء في الإستجابة الذي يتميز به الذين يعانون من الشلل الدماغي وغيره من الإعاقات الحركية الشديدة، لا تعطي اختبارات الذكاء التقليدية تقديرات صحيحة للقدرات الإدراكية لهؤلاء الأفراد. وتُمكّننا تعديلات على اختبارات ستانفورد – بينيه وسلسلة ويكسلر وغيرها من اختبارات الذكاء الفردية بالإضافة إلى إجراء الاختبارات ذات المهمة الواحدة مثل مصفوفات ريفين التدرجية وعدد من اختبارات المفردات المصورة دون وقت محدد من الحصول على معلومات عن الوظائف الإدراكية للأفراد. ويمكن للممتحنين المعاقين جسدياً الإجابة على الفقرات عن طريق هز الرأس أو الإشارة.

يعد كل من اختبار بيبودي للمفردات المصورة (PPVT) ومقياس كولومبيا للنضج العقلي (CMMS) من أكثر الاختبارات المصورة شعبية. يتألف اختبار CMMS من ٩٢ بطاقة تحتوي على مجموعة من الصور يتعين على الممتحن أن يتعرف من بينها على الصورة التي لا تنتمي إلى المجموعة. ويتم تحويل مجموع الدرجات الأولية في اختبار CMMS إلى درجات عمرية إنحرافية تتراوح بين ٥٠ و١٥٠. وعلى الرغم من أن قياس الطبعة الثالثة من هذا الاختبار تم في السبعينيات إلا أن معاملات ثباته تبدو مقبولة. ومع ذلك هنالك تحفظان على اختبار CMMS يتلخصان بما يلي: (١) يمكن للتخمين الصحيح (الحظ) أن يؤدي إلى تضخم في درجات الاختبار و(٢) يقيم الاختبار القدرة على تكوين المفاهيم أكثر من تقييمه للقدرة العقلية العامة (كوفمان، ١٩٧٨).

تتألف الطبعة الثالثة من اختبار بيبودي للمفردات المصورة (PPVT) من ١٧٥ بطاقة تحتوي كل واحدة منها على أربع صور. وتصنف البطاقات في

مجموعات من ١٢ بطاقة يتم ترتيبها تصاعدياً من الأسهل إلى الأصعب حسب عمر المتحَن. ويستغرق الاختبار من ٥ إلى ١٠ دقائق فقط وهو مناسب للاستخدام مع الأفراد بين سنتين ونصف إلى خمسة وثمانون سنة. تشابه طريقة إجراء اختبار PPVT-III طريقة إجراء اختبار ستانفورد – بينيه حيث يتم تحديد البطاقة التي يبدأ بها المتحَن تبعاً لعمره الفعلي. ويتم مع سحب كل بطاقة إعطاء المتحَن كلمة (يقرؤها القائم على الاختبار شفهياً) ويطلب منه الإشارة إلى الرسم الذي يعبر عنها بالشكل الأفضل من الرسومات الموجودة على هذه البطاقة. وفي حال فشل المتحَن في الإجابة على الأسئلة المرتبطة ببطاقة ما يتم سحب البطاقات من المستويات الأدنى حتى يُعطي المتحَن عدداً معيناً من الإجابات الصحيحة على التوالي، ويؤدي هذا إلى انتقال المتحن إلى المستوى الأعلى ويستمر صعوده حتى يفشل في الإجابة على عدد معين من الأسئلة على التوالي. يتم تحويل الدرجات الأولية للاختبار (ويمثلها عدد الإجابات الصحيحة) إلى مقياس إنحرافي لمعدل الذكاء وسطه الحسابي ١٠٠ وإنحرافه المعياري ١٥.

لقد تم عبر السنوات استخدام الطبعات المختلفة من اختبار PPVT مع الأطفال الذين يعانون من العجز النطقي والشلل الدماغي وصعوبات القراءة والتخلف العقلي بالإضافة إلى الأطفال المنعزلين والذين يسهل تشتيت انتباههم. كما تم أثناء تصميم وتقنين الطبعة الأخيرة من الاختبار مع مراعاة إزالة التحيز وجعل عملية إجراء الاختبار دقيقة وعلى درجة جيدة من الكفاءة.

الاختبارات العصبية – النفسية

يتألف الجهاز العصبي البشري من مكونات مركزية وأخرى محيطية. ويتألف الجهاز العصبي المركزي من الدماغ والنخاع الشوكي على حين يتألف الجهاز العصبي المحيطي من الأعصاب والجهاز العصبي التلقائي. وعلى الرغم

من أن هنالك اضطرابات يمكن أن تؤثر على الجهاز العصبي المحيطي، يُعنى علماء النفس والمختصون في الإضطرابات العصبية بشكل رئيسي بالإضطرابات التي تؤثر على الجهاز العصبي المركزي وبخاصة الدماغ.

قد يتأثر الجهاز العصبي المركزي بعدد من الاضطرابات الناجمة عن عوامل كيميائية وفيزيائية قد لا تؤثر جميعها على الوظائف النفسية عند المصابين، أما عندما يصل الضرر الذي لحق بالدماغ إلى درجة تتأثر فيها وظائف الإدراك والإدراك الحسي والعواطف وغيرها من الوظائف النفسية يتم الإشارة إلى الحالة على أنها اضطراب عصبي – نفسي (Neuropsychological disorder). ومن الأعراض العامة لسلوك الأفراد الذين يعانون من اضطرابات عصبية – نفسية نتيجة لإصابة دماغية ما: فرط الحركة والتسرع وسهولة فقدان الانتباه أو التشتت وفقدان التوازن العاطفي وعدم التعاون مع الآخرين. وتظهر أعراض أخرى عند الأفراد الذين يعانون من التلف الدماغي كعجز في الإدراك الحسي – الحركي ومشاكل في التنسيق (التناغم) الحركي وقصر مدى الانتباه واضطرابات في التعرف الإدراكي الحسي على الأشياء، وفي تمييز الأشكال والذاكرة والتعلم وإصدار الأحكام. تنتج الاختلالات الوظيفية التالية عن إصابة مراكز اللغة في الدماغ: مرض صعوبة فهم اللغة المحكية والمكتوبة (aphasia) ومرض صعوبة القراءة (alexia) ومرض صعوبة الكتابة (agraphia)؛ أما تأثر المراكز الحسية في الدماغ فقد يؤدي إلى مشاكل في القدرة على التعرف على الأشياء (agnosia) بينما تؤدي إصابة مراكز الحركة إلى عدم القدرة على القيام بالحركات الإرادية (apraxia) وفقدان القدرة على التناغم الحركي كما أن يمكن أن تؤدي كذلك إلى الشلل. ومن أكثر الأمراض العصبية – النفسية شيوعاً عند كبار السن مرض الزهايمر ومرض الخرف متعدد الاحتشاءات إلا أن الأعراض الجسدية والنفسية قد تنتج عن حالات أخرى كالإصابة الدماغية الرضية (TBI) والإدمان المزمن على الكحول والأورام الدماغية ومرض باركنسون.

وبالرغم من التطور والدقة الهائلين اللذين تتمتع بهما الأساليب العلمية الحديثة في تصوير الدماغ (CAT) أو التصوير الطبقي المحوري وMRI أو التصوير بالرنين المغناطيسي وPET أو التخطيط الطبقي بانبعاث البوسيترون وغيرها من الإجراءات التشخيصية، من الصعب جداً تحديد درجة وموضع وتأثيرات التلف العصبي بشكل كامل باستخدام الطرق اللانفسية وحدها. وقد أدى هذا إلى تصميم اختبارات عصبية - نفسية خاصة تقيس الإحساس والسرعة والقوة الحركية والإدراك الحسي والقدرة على الدمج بين الإدراك الحسي والحركة واللغة والإنتباه والقدرة على التجريد والتوجه وعمل الذاكرة. وتلعب نتائج هذه الاختبارات دوراً في التشخيص السريري للحالات النفسية وفي التخطيط لعلاجها، كما أنها تساهم في تحديد مدى الإعاقة التي أصيب بها الفرد وذلك في حالات طلب تعويضات عن إصابات العمل أو تحديد الراتب التقاعدي لبعض الأفراد. وتستخدم هذه الاختبارات كذلك في تقييم القوى العقلية للأفراد وتحديد مسؤوليتهم القانونية عن قرارتهم. تتوفر عدد من الاختبارات العصبية - النفسية في الأسواق عن طريق دور النشر والتوزيع التي تتعامل بأدوات القياس النفسي كشركة مصادر التقدير النفسي وشركة علم النفس المتحدة ومختبر ريتين العصبي - النفسي والخدمات النفسية الغربية. وتتوفر اختبارات مفردة أو بطاريات كاملة وذلك بهدف توفير الحد الأقصى من المرونة في التعرف على الآثار الإدراكية للأمراض العصبية - النفسية وتحديد مواضعها ومداها ودرجة إستجابتها للعلاج. ويمكن للقارئ الحصول على قائمة بالاختبارات العصبية - النفسية وقائمة أخرى باختبارات الذاكرة عن طريق اختيار الفقرات ٥ و٦ من قائمة برنامج H-7 الموجود على القرص المرفق بهذا الكتاب.

وعلى الرغم من أن استخدام الاختبارات النفسية يساعد في تشخيص الإضطرابات العصبية - النفسية، لا يمكن القول بأن هذه الاختبارات تكفي

وحدها لتشخيص الحالة بالشكل المطلوب إذ يجب أن تترافق هذه الاختبارات مع دراسة مستفيضة لتاريخ الحالة ومراقبة وتقييم سلوك المريض وإجراء عدد من الفحوصات الطبية.

استخدام اختبار WAIS كاختبار عصبي - نفسي

من الممكن التعرف على التغيرات التي تطرأ على الأفراد نتيجة الإضطرابات العصبية – النفسية عن طريق إجراء اختبارات الذكاء مثل WISC-III و-WAIS-R وWAIS-III. إذ تشير الاختلافات الواضحة بين درجات الأفراد في اختبارات الأداء والاختبارات اللغوية بالإضافة إلى تشتت في الدرجات المعيارية للاختبارات الفرعية إلى وجود اضطرابات دماغية رضية كما أنها قد تقود إلى تحديد موضع التلف العصبي. ويعتقد ريتين (١٩٦٦) أن انخفاض معدل الذكاء اللغوي لفرد ما عن معدل ذكائه الأدائي بشكل كبير قد يشير إلى خلل في النصف الأيسرمن الدماغ بينما يشير انخفاض معدل الذكاء الأدائي عن معدل الذكاء اللغوي إلى خلل في نصف الدماغ الأيمن. كما يُلاحظ كذلك أن الأفراد الذين تعرضوا إلى إصابات دماغية غالباً ما يحصلون على درجات أقل في اختبارات المتشابهات والمدى الرقمي والرموز الرقمية من غيرها من الاختبارات الفرعية في اختبار WAIS (ماتارازو، ١٩٧٢). إلا أن هذه إشارات فقط وليست دلائل قاطعة على وجود تلف دماغي.

وقد أدت ضرورة التعرف على تأثيرات التلف والإصابات الدماغية على الوظائف الإدراكية والسلوكية بشكل واضح إلى تعديل على اختبار WISC-R يطلق عليه اسم "اختبار WISC-R كأداة قياس عصبية – نفسية" أو WAIS-R NI. وباستثناء بعض التعديلات الطفيفة كتلك التي طرأت على اختبارات تجميع الأشياء، تم الإبقاء على كافة الاختبارات الفرعية لاختبار WAIS-R، وأضيفت الاختبارات التالية عليها: اختبار المعلومات واختبار المفردات

عن طريق أسئلة الخيارات المتعددة، واختبار الرياضيات عن طريق الورقة والقلم واختبار المتشابهات عن طريق أسئلة الخيارات المتعددة وترتيب الجمل والمدى المكاني ونقل الرموز. وتساعد مقارنة النتائج التي يتم الحصول عليها من الاختبارات الفرعية التقليدية والتي تركز على عمل الذاكرة التذكري مع نتائج الاختبارات الفرعية الإضافية والتي تركز على الذاكرة التعرّفية في التوصل إلى تقدير أفضل للوظائف الإدراكية المتضررة وغير المتضررة. وتساعد في هذا الأمر كذلك مقارنة الطريقة التقليدية للاختبارات القديمة مع الطرق البديلة المعتمدة في إجراء الاختبارات الجديدة. كما يعطينا تحليل أخطاء الممتحَن وتحليل الإستراتيجيات التي اعتمدها في الإجابة معلومات تفيد عمليات التشخيص وإعادة التأهيل.

اختبارات الذاكرة الإدراكية - الحسية

أدى الربط بين التشوهات التي تطرأ على الإدراك الحسي Perception والذاكرة عند الذين تعرضوا إلى إصابات دماغية إلى تطوير اختبارات تشخيصية خاصة كاختبار بندر Bender الحركي البصري واختبار بنتون Bender للذاكرة البصرية وهما اختباران غالباً ما يتم إلحاقهما باختبارات الذكاء التقليدية وغيرها من أدوات التقييم النفسي.

تعرض على الممتحن في اختبار بندر الحركي البصري المستند إلى مبدأ غشتالت تسع بطاقات (٤ × ٦ إنشاً) تحتوي على تصماميم هندسية، ثم يطلب منه إعادة رسمها(انظر الشكل ٨-٤). وتشير التشوهات في الأشكال التي يرسمها الممتحن عن الأصل، إلى مشكلات في الإدراك الحسي، إذ لا يرتكب الأطفال فوق عمر الثمانية والذين يملكون معدل ذكاء طبيعي أكثر من خطأين في نقل الأشكال. وغالباً ما تكون التشوهات التالية إشارات على وجود إصابة دماغية ما: تشوهات في رسم الشكل العام وتدوير التصميم ومشكلات في الدمج

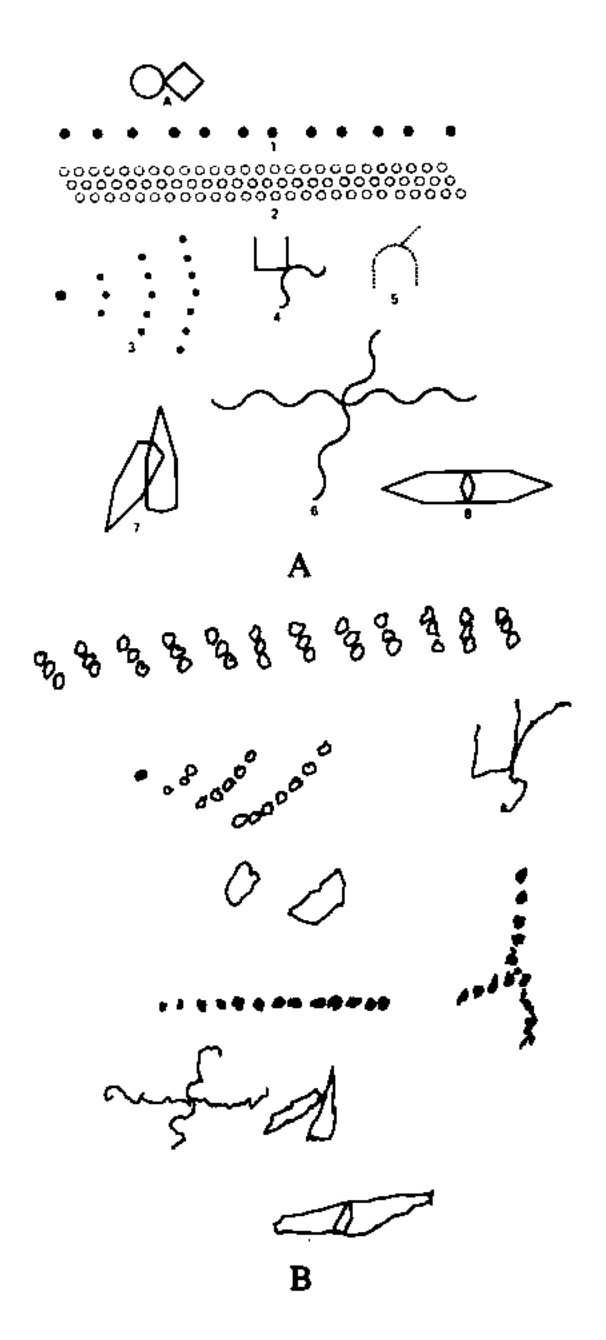
بين التصـمـاميم، ومشـاكل في التناسب الحـجـمي للأشكال، ورسم الأشكال فوق بعـضـهـا، ورسم الأشكال بشكل مـجـزأ، والرسم المتكرر لنفس الشكل حتى بعد غيـاب البطاقـة التي تحمله. ويمكن للقارئ مـراجعـة قـائمـة لاكس (١٩٨٤) التي تصف الأخطاء أو العلامات التشخيصية الثلاثة عشرالمرتبطة باختبار بندر.

يتألف اختبار بنتون للذاكرة البصرية من عشرة تصاميم تقدم كل واحد على حدة للممتحَن ثم يطلب منه رسمها من الذاكرة (على عكس اختبار بندر حيث ينقل الممتحَن الرسومات عن البطاقات الموجودة أمامه). وتفيد رسومات صغيرة في أطراف البطاقات في التأكد من قدرة الممتحَن على تغطية حقل الرؤية بكامله. يتم تقييم اختبار بنتون، كما كان الحال في اختبار بندر، عن طريق تحديد عدد ونوعية الأخطاء التي ارتكبها الممتحَن. وتؤكد الأبحاث التي تمت باستخدام هذا الاختبار ولأكثر من نصف قرن على حساسيته في التعرف على الإصابات الدماغية الرضية ونقص الانتباه وأنواع من الخرف.

اختبارات المسح العصبية

تم إعداد عدد من اختبارات المسح السريعة لاستخدامها كأدوات تقييم أولية قبل استخدام بطاريات الاختبار الكاملة وذلك نظراً لأن عملية الفحص العصبي الكامل تستغرق وقتاً وجهداً كبيرين. ومن الأمثلة على اختبارات المسح هذه: الاختبار العصبي – النفسي الإدراكي المختصر، واختبار المسح العصبي السريع، وأداة مسح بايلي للتطور العصبي للأطفال، واختبار المسح التابع لبطارية لوريا – نبراسكا العصبية – النفسية، واختبار مسح ستروب العصبي – النفسي، واختبار ويسكونسون لتصنيف البطاقات (WCST)، ويعد هذا الأخير أكثر أدوات المسح العصبي – النفسي استخداماً ودراسة[٤].

(٤) بحث كيرت غولدشتاين بشكل مكثف في مدى الارتباط بين الإصابات الدماغية والعجز في قدرة الفرد على القيام بعملية الإستدلال (التفكير) التجريدي بالرغم من عدم حدوث=



الشكل ٨-٤

أشكال من اختبار بندر الحركي البصري المستند إلى مبدأ غشتالت. الشكل (A) هو الشكل الأصلي كما يظهر على البطاقة بينما يحمل الشكل (B) رسم الشكل ذاته من قبل شخص يعاني من تلف دماغي عضوي بسبب إدمان الكحول. (الرسم في الأعلى من اختبار بندر، ١٩٣٨ حقوق الطبع ١٩٣٨ لإتحاد أطباء العظام - النفس. تمت إعادة الطبع بالموافقة. الرسم في الأسفل من لاكس، ١٩٨٤، ص٣٤ حقوق الطبع ١٩٨٤ محفوظة لجون وايلي وأبنائه. تمت إعادة الطبع بالموافقة).

= أي تغيير في قدرة هذا الفرد على التفكير الملموس. وقد طور غولدشتاين وزملاؤه اختبارات غولدشتاين - شيرر للتفكير التجريدي والملموس والتي أنحدرت منها اختبارات مثل اختبار WCST.

يستغرق إجراء اختبار ويسكونسون لتصنيف البطاقات (WCST) المناسب للأفراد بين أعمار السادسة والنصف وثمانين عاماً من ٢٠ إلى ٣٠ دقيقة على الرغم من أن وقت الاختبار بحد ذاته غير محدود، ويعتبر هذا الاختبار حساساً لاضطرابات الفص الجبهي بشكلٍ خاص كما أنه يفيد في تقييم التفكير التجريدي والنزعة إلى التكرار ويساعد في تحديد ما إذا كانت الإصابة الدماغية في الفص الجبهي أم لا. يستخدم الاختبار ٤ بطاقات تحفيزية ورزمة مكونة من ٦٤ بطاقة إستجابة تضم من ١ إلى ٤ رموز (مثلث، أو نجمة، أو صليب، أو دائرة) ملونة بأحد الألوان التالية: أحمر أو أصفر أو أخضر أو أزرق. ويطلب من المتحَن تصنيف بطاقات الاستجابة تحت كل من البطاقات التحفيزية وفقاً لمبدأ مجهول (يتعين على المتحَن التوصل إليه وقد يكون هذا المبدأ مرتبط باللون أو الشكل أو العدد). يتم إعلام المتحَن بصواب أو خطأ إجابته وعند تحقيقه لعشر إجابات صحيحة متتالية يتم تغيير مبدأ التصنيف (من الشكل إلى اللون مثلاً)، وعلى الممتحن أن يتوصل إلى المبدأ الجديد ليصنف على أساسه. يتم منح المتحن درجات تعتمد على عدد المحاولات التي قام بها قبل أن يتوصل إلى عدد معين من الإجابات الصحيحة المتتابعة تبعاً لمبدأ التصنيف المطلوب. يضم دليل المستخدم المعدل للاختبار معلومات عن ثبات وصدق ومعايير الاختبار إلا أن إيغلاند (١٩٨٥) ينصح بتوخي الحذر عند استخدام هذا الاختبار لأغراض سريرية، كما يشكك كل من سنو وماونتين (١٩٩٣) في مدى حساسية اختبار WCST لإصابات الفص الدماغي الجبهي.

ضعف الذاكرة واختباراته

تعتبر مشاكل ضعف الذاكرة طويلة الأمد وقصيرة الأمد دلالات لا على التخلف العقلي فحسب بل على عدد من المشاكل الأخرى كصعوبات التعلم الخاصة، والإصابات النخاعية، والإضطرابات العصبية، ومرض نقص الانتباه، وفرط الحركة (ADHD)، والتقدم في العمر، وحتى الإضطرابات العاطفية.

وغالباً ما يعاني الأفراد الذين تعرضوا إلى إصابات دماغية من إنحدار في قدرات الذاكرة التذكرية recall memory (وبخاصة في حالات التذكر الحر) أكثر من الذاكرة التعرفية recognition memory. ولذلك لا تظهر مشاكل الذاكرة عند هؤلاء المرضى بشكل واضح في الاختبارات التي تقيِّم الذاكرة التعرفية بينما تبدو جلية في اختبارات الذاكرة التذكرية. ونظراً لتأكيد اختبارات الذكاء الفردية كسلسلة ويكسلر على التذكر الحر، تبدو الإعاقة لدى المصابين بالإصابات الدماغية أسوأ مما هي عليه في الواقع[٥].

ويتوفر كم هائل من اختبارات الذاكرة قصيرة الأمد والذاكرة طويلة الأمد ويمكن الحصول على قائمة بهذه الاختبارات عن طريق اختيار الفقرة ٥ من قائمة التشغيل لبرنامج H-7 الموجود على القرص المرافق لهذا الكتاب. كما تتوفر بطاريات اختبار للذاكرة إلى جانب الاختبارات المفردة (أي التي تنشر كل واحد على حدة) واختبارات الذاكرة التي تشكل واحداً من الاختبارات الفرعية ضمن اختبارٍ أشمل. ومن الأمثلة على هذه الاختبارات: مقياس ويكسلر للذاكرة، ومقياس الذاكرة والتعلم، والتقييم الشامل للذاكرة والتعلم، ومقياس تقييم الذاكرة. وتعتبر هذه الاختبارات مصدراً إضافياً للمعلومات التشخيصية وليس بديلاً عن أدوات التقييم الإدراكي الشاملة التي توفرها اختبارات الذكاء الفردية مثل: WAIS-R وWISC-III وK-ABC.

يقيس مقياس ويكسلر للذاكرة (الطبعة الثالثة) WMS-III ذاكرة الأفراد للمحفزات السمعية/اللغوية، والبصرية/اللالغوية، وللأشياء المجردة والملموسة وذلك في نمط التذكر المباشر واللاحق (المتأخر).يتألف الاختبار من ستة

(٥) مما يلاحظ على المصابين بفقدان الذاكرة أن ذاكرتهم للأحداث تكاد تكون شبه معدومة على حين أن ذاكرة المهارات لديهم لا تتأثر بفقدان ذاكرة الأحداث. ففي دراسة أجراها كوهين (١٩٨٤) لوحظ أن فاقدي الذاكرة لا يتذكرون حل أحجية هانوي Hanoi Puzzle (راجع برنامج H-7) في يوم سابق إلا أن سرعتهم عند حل الأحجية في يوم لاحق كانت أسرع مما يشير إلى عدم تأثر ذاكرة المهارات لديهم بفقدان ذاكرة الأحداث.

اختبارات فرعية أساسية وثلاثة ملحقة مصنفة ضمن ثلاثة مجالات، ويستهدف الاختبار الأفراد بين أعمار ١٦ و٨٩ (انظر الجدول ٨-١). ويتم جمع الدرجات المعيارية للاختبارات الفرعية الأساسية والفرعية ثم تحويلها إلى سبعة مؤشرات أولية (هي: الذاكرة العامة، والذاكرة السمعية/اللغوية المباشرة، والذاكرة البصرية/اللالغوية المباشرة، والذاكرة السمعية/اللغوية اللاحقة، والذاكرة البصرية/اللالغوية اللاحقة، والتعرف اللاحق، والذاكرة العاملة) وأربعة مؤشرات وظيفية (حفظ المعلومات في الذاكرة، وإسترجاع المعلومات من الذاكرة، ومنحدر التعلم، والتعلم عن طريق المحاولة الواحدة). تم قياس اختبار WMS-III على عينة طبقية عشوائية مؤلفة من ١٢٥٠ بالغاً. وتم استخدام العينة ذاتها لقياس اختبار WISC-III[٦] مما يسمح بالتعرف عن عجز ومشاكل الذاكرة عند مقارنة معدلات الذكاء الناتجة عن اختبار WISC-III ودرجات اختبار الذاكرة.

الجدول ٨-١

تصنيف الاختبارات الفرعية في مقياس ويكسلر للذاكرة (الطبعة الثالثة) WMS-III

	مجال الذاكرة		
الاختبار الفرعي	سمعي/لغوي	بصري/لالغوي	العمل
أساسي	الذاكرة المنطقية المتشابهات التي تمت المزاوجة بينها على أساس لغوي	الصور العائلية تذكر الوجوه	سلسلة الأرقام والأحرف المدى المكاني
ملحق	قائمة الكلمات	إعادة الإنتاج البصري	السيطرة العقلية

(٦) يستخدم لفظ conformed عند الإشارة إلى اختبارين (للتحصيل والذكاء) تم قياسها على نفس العينة مما يسمح بمقارنة درجات الأفراد في الاختبارين بشكل مفيد لعمليات البحث والتشخيص... إلخ. (المترجم).

ويعد مقياساً الذاكرة والتعلم (TOMAL) والتقييم الشامل للذاكرة والتعلم (WRMAL) من أكثر اختبارات الذاكرة استخداماً مع المراهقين والبالغين. يستغرق كلا الاختبارين ٤٥ دقيقة وتنتج عنهما درجات اختبارات فرعية منفردة أو مركّبة. كما يستغرق إجراء مقياس تقدير الذاكرة (MAS) والذي يستخدم مع المراهقين والبالغين في العمر ١٨ إلى ٩٠ سنة ٤٥ دقيقة كذلك. تقيس الاختبارات الثلاثة المذكورة أعلاه وظائف الذاكرة اللغوية واللالغوية (البصرية) وتعتبر نسب ثباتها مقبولة. ولعل أكثر عناصر دليل المستخدم لاختبار MAS فائدة احتواؤه على مظاهر للمصابين بالخرف وإصابات الرأس المغلقة، وإصابات نصف الدماغ الأيسر وإصابات نصف الدماغ الأيمن.

البطاريات العصبية - النفسية

لا تكفي اختبارات الذكاء العام والاختبارات الخاصة والمحدودة التي تقيس الإدراك الحسي والذاكرة في إجراء تقدير عصبي - نفسي شامل مما يجعل الحاجة ملحة لاستخدام سلاسل من الاختبارات كتلك الموجودة في بطارية هالستيد - ريتين العصبية - النفسية أو بطارية لوريا - نبراسكا العصبية - النفسية، وذلك بهدف الحصول على معلومات تشخيصية دقيقة. وتفيد هذه البطاريات في تقييم العديد من الوظائف العصبية - النفسية والمشاكل الإدراكية الناجمة عن الإصابات الدماغية العضوية.

تتألف بطارية هالستيد - ريتين العصبية - النفسية من مجموعات متعددة من الاختبارات تم تجميعها سويةً بناءً على الفرد المستهدف بعملية القياس (بالغ أو طفل صغير أو طفل أكبر عمراً). وتتألف البطارية الأساسية من الاختبارات الخمسة الأولى المذكورة في الجدول ٨-٢ بالإضافة إلى إجراءات اختبارية أخرى تتضمنها البطارية.

الجدول ٨-٢

اختبارات وإجراءات بطارية هالستيد - ريتين

اختبار الفئات: يحاول الممتحن إيجاد القاعدة التي يمكن على أساسها تصنيف صور تحتوي على أشكال هندسية. ويقيس هذا الاختبار التفكير التجريدي وتشكيل المفاهيم عند الممتحنين.

اختبار الأداء اللمسي: يقوم الممتحن الذي تم وضع عصبة على عينيه بوضع عدد من المكعبات في المكان المناسب لها المحفور في لوح قائم. ويستخدم الممتحن يده الأكثر استخداماً أول الأمر ثم يستخدم يده الأقل استخداماً ثم يستخدم كلتا يديه. يقيس هذا الاختبار القدرات الحركية والحسية - الحركية بالإضافة إلى ذاكرة الأحداث.

اختبار الإدراك - الحسي للأصوات اللغوية: يختار الممتحن من أربعة اختيارات الشكل المكتوب لكلمات لا معنى لها موجودة على شريط مسجل. يقيس الاختبار الإنتباه والتوفيق بين المهارات السمعية - البصرية.

اختبار سيشور للنغمات: يحدد الممتحن ما إذا كانت نغمتان يستمع إليهما متطابقتين أم مختلفتين ويقيس هذا الاختبار الانتباه والإدراك - الحسي السمعي.

اختبار النقر بالأصابع: ينقر الممتحن على أداة تشبه زر التلغراف بأقصى سرعة ممكنة ولمدة ١٠ ثواني متواصلة. يقيس هذا الاختبار سرعة الحركة.

قوة القبضة: يضغط الممتحن على جهاز قياس القوة الميكانيكية (ديناموميتر) بكل قوته، ويمنح عدة محاولات لكل يد. يقيس هذا الاختبار قوة القبضة.

تقفي الأثر (الأجزاء أ و ب): يصل الممتحن بالقلم بين عدد من الأرقام المتسلسلة (الجزء أ) أو عدد من الأرقام والأحرف (الجزء ب) حسب ترتيبها المتسلسل وبالتناوب وذلك ضمن فترة زمنية محددة. يقيس هذا الاختبار القدرة على المسح السريع للمعلومات والمرونة العقلية والسرعة.

التعرف اللمسي على الأشكال: يحاول الممتحن التعرف على عدد من الأشكال البسيطة (مثلث... إلخ) توضع في قبضة يده. يقيس هذا الاختبار القدرة الإدراكية - الحسية للحواس.

الامتحان الإدراكي – الحسي للحواس: ينفذ الممتحن مهام حسية بسيطة (كالتعرف على الإصبع الذي تم لمسه أو على الأذن التي استقبلت صوتاً بسيطاً). يقيس هذا الاختبار القدرة الإدراكية – الحسية للحواس.

اختبار التحري عن الإصابة بحبس النطق (الافازيا): تتضمن مهام هذا الاختبار تسمية الأشياء المصورة (مثال: شوكة) وتكرار عبارات قصيرة والنسخ. يقيس هذا الاختبار القدرات اللغوية الإستقبالية والتعبيرية.

الاختبارات الملحقة: يمكن إجراء أي من الاختبارات التالية بالإضافة إلى الاختبارات الأساسية المذكورة أعلاه: WAIS-R وWRAT-R وMMPI واختبار ويكسلر للذاكرة واختبار راي السمعي اللغوي للتعلم.

من كتاب روبرت ج. غريغوري، الاختبارات النفسية: التاريخ والمبادئ والتطبيق. حقوق الطبع: ١٩٩٢ لألين وبيكون. تمت إعادة الطباعة بتصرف مع الموافقة.

تقيس هذه الاختبارات الوظائف التالية: التفكير التجريدي، والانتباه، والإدراك –الحسي السمعي، والتوافق السمعي– البصري، وذاكرة الأحداث والقدرات الحركية والحسية – الحركية، والمرونة العقلية، والسرعة الحركية، والقدرة اللغوية الإستقبالية، والقدرة على المسح، والقدرة الإدراكية – الحسية للحواس. ويعد اختبارا الفئات وتقفي الأثر الاختبارين الأكثر تعقيداً في بطارية هالستيد – ريتين، إذ يتوجب على الممتحَن في اختبار الفئات استنتاج المبدأ العام الذي يحكم المعلومات المقدمة له على شرائح متتابعة، بينما يتوجب على الممتحَن الوصل بين عدد من الأرقام والأحرف المكتوبة ضمن دوائر حسب ترتيبها المتسلسل وبالتناوب (١ ثم أ، ٢ ثم ب وهكذا). وتستغرق البطارية بأكملها ٦ إلى ٨ ساعات.

تشابه بطارية لوريا – نبراسكا العصبية – النفسية بطارية هالستيد – ريتين من حيث توفيرها لتقدير عصبي – نفسي شامل للتلف الدماغي العضوي. إن البطارية مصممة للأفراد في سن الخامسة عشرة وما فوق ويمكن وضع درجاتها

نسبةً إلى ١٢ مؤشر سريري و٨ مؤشرات موضعية، و٥ مؤشرات خلاصية، ومؤشرين خياريين، و٢٨ مؤشر عاملي يعكسون عدداً من الوظائف الإدراكية والحسية الخاصة. ومن ضمن المتغيرات التي تقيسها بطارية لوريا – نبراسكا: المُسيطر الدماغي، والوظائف اللمسية، والحركية، والبصرية؛ والإدراك – الحسي ثم إعادة إنتاج النبرات والأنغام؛ واللغة التعبيرية والإستقبالية؛ والقراءة والكتابة والحساب؛ وتكوين المفاهيم. تستغرق بطارية لوريا – نبراسكا المؤلفة من ٢٦٩ إلى ٢٧٩ فقرة من ساعة ونصف إلى ساعتين ونصف، وهذا ثلث الوقت الذي تستغرقه بطارية هالستيد – ريتين. يتم منح ما أقصاه ثلاث درجات لكل فقرة اختبارية وذلك بناءً على درجة دلالتها (دلالة قوية أو ضعيفة أو "لا دلالة") على وجود اضطراب في الدماغ. وتعد بطارية لوريا – نبراسكا (جنباً إلى جنب مع بطارية هالستيد – ريتين) من أكثر البطاريات موضع البحث في المجال العصبي – النفسي، وتركز التحفظات والإنتقادات التي وجهت لهذه البطارية على اعتمادها الكبير على المهارات اللغوية وكون ثباتها موضع شك وعدم قدرتها على التعرف على حالات حبس النطق (الأفازيا) وغيرها من الإضطرابات الدماغية بشكل جيد.

تتوفر برامج على حاسب آلي لإجراء عناصر من بطاريتي لوريا – نبراسكا وهالستيد – ريتين بالإضافة إلى عدد آخر من الاختبارات العصبية – النفسية (كاختبار الفئات واختبار ويسكونسون لتصنيف البطاقات). وعلى حين يمكن إجراء هذه الاختبارات إما عن طريق الحاسب أو عن طريق الأفراد، لا يمكن إجراء بعض البطاريات (مثل اختبار MicroCog واختبار تقدير الوظائف الإدراكية) إلا عن طريق الحاسب. يتألف اختبار MicroCog من ١٨ اختباراً فرعياً ويصلح لتقييم البالغين الذين يعانون من اضطرابات طفيفة إلى متوسطة في الوظائف الإدراكية. يستغرق الاختبار من ٥٠ إلى ٦٠ دقيقة، ويمكن الحصول من اختباراته الفرعية على درجات في ٩ مجالات للوظائف الإدراكية هي:

الانتباه/السيطرة العقلية، والذاكرة، والإستدلال/الحساب، والمعالجة المكانية للمعلومات، وزمن رد الفعل وصحة معالجة المعلومات، وسرعة معالجة المعلومات والعمل الإدراكي والبراعة الإدراكية.

الصعوبات التعليمية

غالباً ما كانت الصعوبات في القراءة والكتابة والحساب والتهجئة (وغيرها من المشاكل الأكاديمية) تُعزى إلى التخلف العقلي أو الإعاقات الجسدية أو المشاكل العاطفية الشديدة أو نقص الدافع والحافز. وفي كثير من الحالات تُستبعد جميع الأسباب الأنفة الذكر وتبقى مشكلة الضعف الأكاديمي موجودة كحالات عسر القراءة (Dyslexia) وعسر الحساب (Dyscalculia). ويشار إلى هذه الحالات اليوم على أنها صعوبات تعليمية خاصة أو صعوبات تعليمية. ومن الجدير بالذكر أن هذه الحالات قد تصيب أفراداً على درجة عالية من الذكاء، وعلى عكس الذين يعانون من التخلف العقلي تنخفض معدلات التحصيل بشكل ملحوظ عند أصحاب الصعوبات التعليمية عن قدراتهم الإدراكية.

يحيط كثير من الجدل والاختلاف بتعريف الصعوبات التعليمية ويصعب تحديد العوامل التي تؤدي إلى هذه الصعوبات (عصبية – نفسية، إرتقائية، تجاربية أو عدد من العوامل المجتمعة) إلا أنه يوجد قبول عام لتعريف الفقرات ١٠١–٤٧٦ من القانون لمرسوم الأفراد الذين يعانون من الصعوبات التعليمية لعام ١٩٩٠ (IDEA) والذي ينص على التالي:

يستخدم مصطلح "الأطفال ذوي الصعوبات التعليمية الخاصة" مع الأفراد الذين يعانون من اضطراب معين في واحدة (أو أكثر) من العمليات الإدراكية المطلوبة لفهم اللغة (المكتوبة أو المحكية) مما يؤدي إلى عدم القدرة الكاملة على الإنصات، أو التفكير، أو الكلام، أو القراءة، أو الكتابة، أو التهجئة، أو إجراء العمليات الحسابية. وتشمل الاضطرابات هذه الإعاقات الإدراكية –

الحسية، والإصابات الدماغية، والاضطراب الدماغي الوظائفي البسيط، وعسر القراءة، والأفازيا الإرتقائية. ولا يشمل هذه التعريف الأطفال الذين يعانون من الصعوبات التعليمية الناتجة عن الإعاقات السمعية والبصرية والحركية أو التخلف العقلي أو الاضطرابات العاطفية أو القصور في المزايا البيئية والاجتماعية والحضارية والإقتصادية المتاحة لهم.

يتعرض الأطفال الذين يعانون من الصعوبات التعليمية إلى مشاكل في الإدراك – الحسي والتناغم الحركي والسلوك (كفقد الانتباه وفرط الحركة والتسرع والنزعة إلى التكرار) بالإضافة إلى صعوبات القراءة واللغة والذاكرة والتفكير. ويمكن للمدرسين التعرف على هذه الحالات عن طريق استخدام سلالم التقدير (مثل: الإجراء التقديري للصعوبات التعليمية وسلالم تقدير الطلاب) والاختبارات (مثل: اختبار ماكارثي للمسح واختبار سلينغرلاند للمسح) مع وضع الأطفال تحت الملاحظة الدقيقة. ويجب التذكير هنا أن إجراء بطارية اختبار نفسية شاملة للطفل لا يمكن أن يتم سوى عن طريق المختصين في علم النفس.

يعد التشخيص الفعّال والتخطيط العلاجي لصعوبات التعلّم مشروعاً متعدداً التخصصات يتضمن كلاً من المدرس أو المدرسة المسؤول عن الطفل بشكل مباشر والإخصائي المختص بالصعوبة التعليمية الخاصة التي يتوقع أن الطفل يعاني منها بالإضافة إلى أخصائي في إجراءات القياس النفسي التشخيصية. تنص الفقرات ٩٤-١٤٢ من القانون العام لمرسوم تعليم كافة الأطفال المعاقين لعام ١٩٧٥ على أنه لا يتم تشخيص أي حالة من الصعوبات التعليمية إلا في حال وجود فرق واضح بين مستوى التحصيل والقدرة الإدراكية عند طفل ما في واحدة أو أكثر من النواحي التالية:

- التعبير الشفهي
- الإستيعاب السمعي
- التعبير الكتابي
- مهارات القراءة الأساسية
- الإستيعاب النصي
- الحسابات الرياضية
- الاستدلال الرياضي.

ويتم بناءً على التشـخيص إعداد خطة تعليمية خاصة individual (IEP) education plan تضم أهداف قصيرة وبعيدة الأمد للطفل بالإضافة إلى طرق تحقيق هذه الأهداف.

تختلف معايير تقديم الخدمات التعليمية للأطفال ذوي الصعوبات التعليمية من ولاية إلى أخرى، إلا أن القاعدة العامة تنص على أنه لا يمكن اعتبار طفل ما طفلاً ذو صعوبات تعليمية إلا في حال كون درجاته في اختبار تحصيلي قياسي أقل بدرجة إنحراف معياري كاملة من درجاته في اختبار ذكاء تم معايرته بالتوازي مع اختبار التحصيل. وتعتبر اختبارات الذكاء التالية: SB-IV وWPPSI-R وWISC-III وK-ABC مناسبة للاستخدام مع الاختبارات التحصيلية التالية: اختبار بيبودي الفردي للتحصيل (النسخة المراجعة) واختبار كوفمان للتحصيل التعليمي واختبارات ويكسلر الفردية للتحصيل. إلا أن الخيار الأمثل والأكثر استخداماً في هذه الحالات هو استخدام بطارية وودكوك -جونسون التعليمية - النفسية والتي تتألف من بطارية للذكاء (بطارية WJ-R للقدرة الإدراكية) وبطارية للتحصيل (بطارية WJ-R للتحصيل) تمت معايرتها بالتوازي مع بطارية الذكاء. ويعتمد في تشخيص بعض الحالات على اختبارات عصبية - نفسية خاصة أواختبارات إرتقائية أو اختبارات للشخصية.

الخلاصة

استخدمت اختبارات القدرة الإدراكية بشكل واسع في القرن العشرين وذلك في المجالات السريرية والبحث العلمي. وقد ساهمت التحريات في التطور العقلي عند الرضع والأطفال كتلك التي قام بها أرنولد جيزل وزملاؤه من جامعة ييل في العشرينيات والثلاثينيات في تحديد الطرق ووضع التوصيات لعملية البحث العلمي الذي يعتمد على الأطفال.

كما ساهمت جداول جيزل الإرتقائية ومقياس بايلي لنمو الرضع واختبار دنفر للمسح الإرتقائي وغيرها من الاختبارات القياسية في تدعيم المعرفة العلمية في مجال التطور العقلي واضطراباته عند الأطفال.

وعلى الرغم من الكفاءة والمثابرة لمصممي اختبارات ذكاء الرضع، تبقى هذه الاختبارات أدوات غير ثابتة بالشكل المطلوب وغير قادرة على التنبؤ الدقيق بالتطور والأداء الإدراكي للأطفال في المستقبل، إذ تساهم عوامل كضعف أو عدم الانتباه والتركيز عند الرضع والأطفال دون سن المدرسة والطبيعة الحسية – الحركية لمهام الاختبارات في إنخفاض الإرتباط بين نتائج الاختبارات للأطفال بين أعمار ٢ و٣ سنوات ونتائجهم لاحقاً بعد دخولهم المدرسة.

تحدد المعايير التي ينص عليها مرسوم الأفراد الذين يعانون من الصعوبات التعليمية لعام ١٩٩٠ (IDEA) عمل كافة الاختبارات المصممة للتعرف على وتقييم التأخر الإرتقائي عند الأطفال. ومن الأمثلة على الاختبارات التي تم تصميمها بناءً على المناطق الخمسة التي حددها المرسوم: الاختبارالمسّاح لتقييم الأطفال دون سن المدرسة FirstSTEP ومظاهر AGS للمسح المبكر. كما يتقيد مقياس ماكارثي لقدرات الأطفال وهو اختبار قديم بعض الشيء إلا أنه لا يزال يستخدم بشكل كبير بتوصيات IDEA.

ساهم احتواء مقياس ويكسلر للذكاء على مقياسين منفصلين للقدرات اللغوية والأدائية في إمكانية استخدام هذا الاختبار مع الأفراد الذين يعانون من العجز البصري والسمعي. إلا أن هذا لم يمنع من تصميم اختبارات خاصة بأصحاب الإعاقات الجسدية كالمقياس اللمسي لذكاء البالغين فاقدي البصر واختبار الاستعداد التعليمي للمكفوفين واللذان يستهدفان أصحاب الإعاقات البصرية، واختبار Leiter-R واختبار هيسكي – نبراسكا لقابلية التعلم لأصحاب الإعاقات السمعية، واختبار بيبودي للمفردات المصورة ومقياس كولومبيا للنضج العقلي لذوي الإعاقات العظمية والحركية.

يؤثر التلف الدماغي الناتج عن الإلتهابات والرضوض والمواد الكيميائية وغيرها على عواطف الأفراد ودوافعهم وإدراكهم – الحسي وقدراتهم الإدراكية بالإضافة إلى أنه يسبب الألم والإعاقة الجسدية. ويستخدم الأطباء والأخصائيون النفسيون عدداً كبيراً من الطرق والإجراءات بغرض فهم أسباب ونتائج الإضطرابات الدماغية ثم اقتراح توصيات لمعالجتها. تدخل في عمليات التشخيص والعلاج هذه إجراءات عديدة كوضع المريض تحت الملاحظة وجمع المعلومات من أفراد عائلته ومنه شخصياً وإجراء اختبارات عصبية ونفسية للمريض بالإضافة إلى تصوير دماغه بعدة طرق.

يساعد أحد اختبارات مقياس ويكسلر لذكاء البالغين وهو اختبار WISC-R كأداة عصبية – نفسية في تحديد شدة الإضطراب العصبي – النفسي عند الأفراد. وتتوفر كذلك عدة اختبارات عصبية – نفسية خاصة كاختبار بندر الحركي البصري على مبدأ غشتالت واختبار بنتون للذاكرة البصرية اللذان يقيسان الذاكرةالإدراكية – الحسية، واختبار ويسكونسون لتصنيف البطاقات والذي يعتبر أداة مسح سريعة، ومقياس ويكسلر للذاكرة الذي يقيس الذاكرة قصيرة الأمد وطويلة الأمد وغيره من اختبارات الذاكرةكمقياس الذاكرة والتعلم

والتقييم الشامل للذاكرة والتعلم ومقياس تقييم الذاكرة. وينصح عند الرغبة في إجراء تقييم شامل للاضطرابات العصبية – النفسية عند الأفراد باستخدام بطارية كاملة من الاختبارات كبطارية هالستيد – ريتين العصبية – النفسية أو بطارية لوريا – نبراسكا العصبية – النفسية. وعلى حين يمكن إجراء العديد من الاختبارات العصبية – النفسية إما عن طريق الحاسب أو عن طريق الأفراد، لا يمكن إجراء بعض البطاريات كاختبار MicroCog، وهو أداة مسح إدراكية لتقييم البالغين الذين يعانون من اضطرابات طفيفة إلى متوسطة في الوظائف الإدراكية، إلا عن طريق الحاسب.

تُعرّف الإضطرابات التعليمية الخاصة على أنها صعوبات في القراءة والكتابة والحساب والتهجئة وغيرها من المشاكل الأكاديمية التي لا تعزى أسبابها إلى التخلف العقلي أو الإعاقات الجسدية أو الاضطرابات العاطفية أو فقر البيئة المحيطة بالطفل. وينص القانون الفيدرالي على ضرورة تشخيص حالات الأطفال ذوي الصعوبات التعليمية من قبل مختصين ثم وضع خطة تعليمية خاصة لهؤلاء الأطفال. ويعد إنخفاض درجات طفل ما في اختبار تحصيلي قياسي بشكل كبير عن درجاته في اختبار ذكاء تمت معايرته بالتوازي مع اختبار للتحصيل المؤشر النفسي – التشخيصي الأكثر قبولاً في معظم الولايات. وتستخدم اختبارات WISC-III واختبارات ويكسلر الفردية للتحصيل واختبار كوفمان للتحصيل التعليمي بالإضافة إلى اختبارات القدرات الإدراكية والتحصيل من بطارية وودكوك – جونسون التعليمية – النفسية في تحديد الصعوبات العليمية الخاصة عند الأطفال.

مصطلحات للمراجعة

راجع معاني المصطلحات المدرجة أدناه والتي تم استخدامها في هذا الفصل. الرجاء مراجعة فهرس التعريفات أو المعجم في حال وجود أي التباس حول معاني المصطلحات.

السلوك التكيفي الاختبارات المعايرة بالتوازي

مرض صعوبة القدرة على الكلام (Aphasia) العمر التطوري (الارتقائي)

مرض صعوبة القراءة (Alexia) المعدل الارتقائي

مرض صعوبة الكتابة (Agraphia) الإعاقة (Disability)

صعوبة التعرف على الأشياء (Agnosia) صعوبة الحساب (Dyscalculia)

صعوبة القيام بالحركات الإرادية(Apraxia) صعوبة القراءة والكتابة(Dyslexia)

"في خطر" ذاكرة الأحداث

الجهاز العصبي المركزي الصعوبة

اللمسي الاضطراب العصبي - النفسي

الصعوبة (Handicap) العجز (Impairment)

الخطة التعليمية الفردية (IEP) الجهاز العصبي المحيطي

الصعوبات التعليمية ذاكرة المهارات

الصعوبات التعليمية الخاصة

المراجع المقترحة

Albert, M. S. (1996). Neuropsychological testing. In J. Sadavoy, L. W. Lazarus, L. F. Jarvik, & G. T. Grossberg (Eds.), *Comprehensive review of geriatric psychiatry: II* (2nd ed., pp. 325–350). Washington, DC: American Psychiatric Association.

Detterman, D. K., & Thompson, L. A. (1997). What is so special about special education? *American Psychologist, 52,* 1082–1090.

Goodman, J. F. (1990). Infant intelligence: Do we, can we, should we assess it? In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence and achievement* (pp. 183–208). New York: Guilford Press.

Johnson, J. H., & Goldman, J. (1993). Approaches to developmental assessment. *Advances in Clinical Child Psychology, 15,* 243–274.

Pennington, B. F. (1991). *Diagnosing learning disorders: A neuropsychological framework.* New York: Guilford Press.

Reynolds, C. R. (1990). Conceptual and technical problems in learning disability diagnosis. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence and achievement* (pp. 571–592). New York: Guilford Press.

Sullivan, P. M., & Burley, S. K. (1990). Mental testing of the hearing-impaired child. In C. R. Reynolds & R. W. Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence and achievement* (pp. 761–788). New York: Guilford Press.

Taylor, H. G., & Fletcher, J. M. (1990). Neuropsychological assessment of children. In G. Goldstein & M. Hersen (Eds.), *Handbook of psychological assessment* (2nd ed., pp. 228–255). New York: Pergamon.

الملحق أ

برامج الحاسب المستخدمة في الاختبارات والامتحانات[1]

المقدمة

يعتمد تشغيل البرامج الموجودة على هذا القرص المرن على qbasic أو basica يمكن تحميل qbasic من القرص، ويمكن تشغيل البرامج من القرص المرن أو من القرص الصلب (وغالباً ما تكون البرامج أسرع عند تشغيلها على القرص الصلب). ينصح باتباع التعليمات التالية لنسخ الملفات من القرص المرن إلى القرص الصلب:

الحد الأدنى لمتطلبات نظام التشغيل

- IBM-PC أو معادلاته من الأجهزة التي تضم نسخة DOS 3.3 أو أعلى.
- محرك قرص مرن ٣,٥ إنش.

كيفية تثبيت الملفات على الحاسب

ننصحك بنسخ الملفات على القرص الصلب وذلك بتشغيل برنامج التثبيت. ولتثبيت الملفات، إتبع الخطوات التالية:

١- أدخل القرص في محرك القرص المرن.

٢- إطبع A:\INSTALL عند ظهور أمر DOS، اضغط مفتاح الإدخال. ستظهر شاشة تثبيت البرنامج، اضغط مفتاح الإدخال للتتابع.

٣- إن الفهرس الذي يظهر تلقائياً لتخزين الملفات هو C:\TESTS، ويمكن تغيير هذا الموقع إذا شئت. نفذ التعليمات التي تظهر على الشاشة.

---

(١) يعتمد تشغيل هذه البرامج على qbasic وbasica. يتوفر على القرص المرن الملف التنفيذي qbasic.exe والذي وفرته لنا شركة مايكروسوفت. جميع الحقوق محفوظة لشركة مايكروسوفت.

٤- سيقوم برنامج التثبيت بنسخ كافة الملفات بما فيها برنامج qbasic إلى قرصك الصلب في فهرس C:\TEST أو أي فهرس آخر يحدده المستخدم.

كيفية استخدام البرامج

عند تشغيل البرامج من القرص المرن قم بعمل التالي:

١- أدخل القرص المرن في مشغل الأقراص المرنة في جهاز الحاسب.

٢- إطبع A: في DOS عند ظهور أمر DOS

٣- اطبع menu أو prog في سطر < A:\ ثم اضغط مفتاح الإدخال. الأمر menu سوف يدخلك إلى قائمة البرامج والتي تمكنك من اختيار البرنامج الذي تريد تشغيله. يتبع الأمر prog السؤال: "category and number of program? والتي يتوجب عليك تحديد فئة لها من (A إلى H) ورقم البرنامج الذي ترغب بتشغيله مثلاً (prog.al).

وبإمكانك كذلك تشغيل البرنامج وأنت في Windows من قائمة File في program manager أو في قائمة start اختر الأمر Run اطبع a: menu أو a: prog في حقل التعليمات ثم اضغط الزر Ok.

إذا قمت بتثبيت البرنامج على قرص الحاسب الصلب قم بعمل التالي:

١- في Dos اطبع cd c:\tests (أو الفهرس الذي اختاره المستخدم) واضغط على مفتاح الإدخال.

٢- اطبع menu أو prog في سطر <c:\ tests وأضغط مفتاح الإدخال. الأمر menu سوف يدخلك على قائمة البرامج والتي بإمكانك اختيار البرنامج الذي تريد تشغيله منها. يتبع الأمر prog السؤال: "category and number of pro- gram? مما يتطلب منك تحديد فئة البرنامج (من A إلى H)، بإمكانك الخروج من أي برنامج قيد التشغيل بضغط المفتاح F1.

إن مخرجات بعض البرامج سوف تظهر على الشاشة أمامك و/أو في ملف مخرجات على القرص يحمل اسم results. اتبع التعليمات لكل برنامج بحذر ولن تحصل معك مشاكل تذكر. وتذكر دائماً أن الامتحانات المدرجة في بعض البرامج يقصد بها أن تكون تمارين توضيحية أو للعرض فقط أكثر مما هي اختبار جدي بهدف التعليم.

لا تتوفر معايير تمثيلية لعينات الاختبارات الموجودة، ولذلك فمن الأفضل استعراض النتائج على أنها نتائج مقترحة وليست نهائية أو حاسمة.

مساعدة المستخدم

إذا احتجت لمساعدة مبدئية في عملية التثبيت أو اضطررت لاستبدال قرص مرن تالف، يرجى الاتصال بقسم دعم المنتجات على الرقم ٦٧٥٣-٨٥٠ (٢١٢) خلال أيام الأسبوع من التاسعة صباحاً وحتى الرابعة بعد الزوال بتوقيت شرق أمريكا.

للقيام بطلبيات إضافية أو لطلب معلومات إضافية عن منتجات Wiley، يرجى الإتصال بـ ٥٩٤٥ ٢٢٥ (٨٠٠).

معلومات عن البرنامج

فيما يلي وصف مختصر قد يساعدك في التعرف على البرامج:

برامج الفئة A: الطرق الأساسية في الإحصاء

١- التوزيعات التكرارية و الرسوم البيانية المرتبطة بها: يقوم هذا البرنامج بإنشاء التوزيعات التكرارية للدرجات وإعداد المدرجات التكرارية والمضلعات لمجموعة من الدرجات. يقوم المستخدم بتحديد عدد المجالات التي يرغب بها وفيما إذا كان يحتاج لمدرج أولمضلع التكرار. وتظهر على الشاشة، بالإضافة إلى المجالات والتكرارات المقابلة لها، حدود الدرجات العليا والدنيا.

٢- الإحصاء الوصفي: بناءً على مجموعة من الدرجات غير المصنفة والتي لا تتجاوز ١٠٠ درجة، يقوم هذا البرنامج بحساب وطبع الحدود العليا والدنيا للدرجات، ومدى الدرجات، والوسط الحسابي، والوسيط، والتباين، والإنحراف المعياري، ومؤشرات الإلتواء والتفرطح، والدرجات الخام والدرجات z الموافقة لها.

٣- برنامج رسم مخطط التبعثر وخط الإنحدار: يقوم هذا البرنامج برسم مخطط التبعثر وخط الإنحدار لـ yعلى xولـ xعلى yلجملة من معدلات الدرجات. كما يقوم بحساب وطباعة القيم العددية للمنحدر ويتوقع مجال التقاطع لتنبؤ X من Y وY من X ومعامل الإرتباط المولد للعزوم (r).

٤- الإنحدار والإرتباط الخطي البسيط: يقوم هذا البرنامج بحساب معامل الإرتباط المولد للعزوم (r) بين X وY، كما يقوم بحساب معادلة الإنحدار الخطي للتنبؤ بـ Y من X والأوساط الحسابية والإنحرافات المعيارية لـ X وY والخطأ المعياري للتقدير وقيمة z لاختبار للدلالة الإحصائية لمعامل الارتباط لـ ٣ - ١٠٠ زوج من قيم X وY. كما يمكن تحديد عدد قيم الأخطاء السلبية والإيجابية غير الحقيقية وعدد حالات القبول والرفض للممتحنين الذين وقعت درجاتهم دون الحد الفاصل للدرجات (أو دون مستوى الأداء المقبول).

٥- تحليل الإنحدار المتعدد: يقوم هذا البرنامج بحساب أثقال الإنحدار القياسي وغير القياسي ومعامل الارتباط المتعدد (R) والأخطاء المعيارية لأوزان الإنحدار وقيم التحديد الدلالة الإحصائية لأوزان الانحدار لتحليل إنحدار خطي بالنسبة لمتغير مستقل واحد أو لمتغيرين مستقلين أو لثلاثة متغيرات مستقلة. والمتغير الأخير هو متغير تابع. المدخلات هي الأوساط الحسابية والإنحرافات المعيارية والإرتباطات الداخلية بين المتغيرات.

برامج الفئة B: وضع وإجراء وتصحيح الاختبارات

١- بناء الاختبارات الموضوعية: يسمح هذا البرنامج للمستخدم بالقيام بما يلي:

- بناء اختبارات موضوعية للشخصية
- استعراض ومراجعة الاختبار الذي تم بناؤه
- وضع مفتاح الإجابات للاختبار
- استعراض ومراجعة مفتاح إجابات الاختبار
- طباعة الاختبار
- طباعة مفتاح إجابات الاختبار
- تغيير كلمة سر البرنامج
- إدخال سلسلة من أرقام التعريف (الأرقام الامتحانية التي يحملها كل من يتقدم للاختبار) للاحتفاظ بسجل لمن يتم اختباره ولضمان أن يتم اختبار كل فرد مرة واحدة فقط.

إن كلمة السر الأولية للبرنامج هي Makeit. يتم تخزين أرقام التعريف من خلال الاختيار الأخير أعلاه، كما يتم مسح هذه الأرقام بعد إنتهاء الاختبار. يمكن تخزين العديد من الاختبارات ومفاتيح الإجابات المختلفة في ملفات مناسبة على القرص المرن الحاوي للبرنامج أو على قرص مرافق، ويتم تخصيص ملف خاص مشفر لكل اختبار ومفتاح إجاباته. ويمكن تنقيح أو تغيير عنوان وتعليمات وفقرات الاختبار ومفتاح إجاباته بعد الإنتهاء من إجراء الاختبار.

٢- إجراء وتصحيح الاختبارات الموضوعية: يمكن استخدام هذا البرنامج، وهو برنامج مرافق لبرنامج BI، في إجراء وتصحيح أي نوع من الاختبارات الموضوعية (اختبارات الصواب والخطأ، اختبارات الاختيارات المتعددة،

اختبارات المطابقة، اختبارات الترتيب، والاختبارات ذات الإجابات القصيرة... إلخ). وتسمح خيارات البرنامج للمستخدم بالقيام التالي:

- إجراء الاختبار.
- وضع درجات للاختبارات.
- استعراض بنود الاختبار والإجابات.
- طباعة بنود الإجابات ومعدل الدرجات الكلي.

على المستخدم أن يقوم بإدخال كلمة المرور takeit (وهي كلمة المرور التي تأتي مع البرنامج والتي يمكن تغييرها لاحقاً) ورقم الاختبار ورقمه التعريفي. ويعطينا الخيار الأخير أعلاه ورقة مطبوعة تضم رقم الاختبار ورقم المستخدم التعريفي وجدول من ستة أعمدة يدرج العمود الأول رقم السؤال أو الفقرة الاختبارية، ويدرج العمود الثاني أول إجابة يختارها الممتحن لكل سؤال (إذ قد يعود الممتحن ويغير هذه الإجابة)، أما العمود الثالث فيدرج الزمن بالثواني الذي احتاجه الممتحن للتوصل إلى الإجابة، وفي العمود الرابع تُدرج الإجابة النهائية للسؤال (بغض النظر عن عدد المرات التي غير فيها الممتحن رأيه واختياره للإجابة). يدرج العمود الخامس الزمن (بالثواني) الذي احتاجه الممتحن لاختيار إجابته النهائية، أما العمود السادس فيدرج أحد الحرفين R أو W إشارة لكون الإجابة صحيحة R أو خاطئة W. يمكن للممتحن أن يقوم بإجراء الاختبار واستعراض إجاباته وتصحيح الاختبار، ومن ثم طباعة الإجابات والزمن المستغرق ومجموع درجات الاختباروذلك بهدف العودة إلى هذه النتائج لاحقاً أو الاحتفاظ بالنتائج كسجل دائم.

٣- التصحيح من أثر التخمين في الاختبارات الموضوعية: يقوم هذا البرنامج القصير بتحويل الدرجات الكلية للاختبارات الموضوعية إلى درجات مصححة

من أثر التخمين. يقوم المستخدم بإدخال عدد الفقرات الكلي للاختبار وعدد الفقرات الصحيحية والخاطئة وعدد الخيارات (البدائل). وتتضمن ورقة النتائج المطبوعة (ملف النتائج) معدل الدرجات قبل التصحيح ومعدل الدرجات المصحح باستخدام معادلة التصحيح من أثر التخمين التقليدية ومعادلة Little ومعادلة Reid (راجع أيكين، ١٩٨٧).

٤- إجراء وتصحيح فقرات المطابقة وإعادة الترتيب: يمكن إجراء فقرات المطابقة أو إعادة الترتيب والتي تتألف كحد أقصى من ١٥-٢٠ جزء باستخدام هذا البرنامج. يتم إدخال أعمدة تضم جذور الأسئلة وأعمدة تضم الإجابات لفقرات المطابقة من خلال الملفين match1 أو match2 أما الإجابات الصحيحة فيتم إدخالها من الملف match3 وعلى الرغم من أنه تم تخزين معطيات عينة لسؤال مطابقة، يمكن للمستخدم تغيير هذه المعطيات بسهولة وإعداد الفقرات التي يريدها. يتم إدخال معطيات فقرات إعادة الترتيب من ملفات rank1 وrank2 وعلى الرغم من أنه تم تخزين معطيات عينة لسؤال إعادة ترتيب، يمكن للمستخدم تغيير هذه المعطيات بسهولة وإعداد فقرات خاصة به. تتألف نتائج فقرات المطابقة من عدد ونسبة المطابقات الصحيحة، أما نتائج فقرات إعادة الترتيب فتتألف من معدلي درجات تم حسابهما بواسطة المعادلات (3.1a) و(3.1b) في الصفحة ٥٩ من أيكين (a١٩٨٧).

٥- وضع درجات ترتيب الفقرات: يقوم هذا البرنامج بإعطاء الدرجات لفقرات الترتيب أو إعادة الترتيب باستخدام طريقة أيكين، ١٩٧٠، يقوم المستخدم بتحديد عدد الفئات وإدراج التصنيفات الصحيحة (المبوبة) وتصنيفات الممتحن لكل فئة. يمكن طباعة الدرجات إما على شكل أعداد عشرية أو أعداد صحيحة مقربة.

٦- وضع الدرجات باستخدام طريقة كاجوري المعدلة:يعطي هذا البرنامج باستخدام تعديل على طريقة كاجوري (أيكين، ١٩٨٣) التوزيع التكراري لدرجات ومجال درجات مجموعة من الطلاب. ويتوجب على المستخدم تحديد عدد الاختبارات أو معدلات الدرجات والحدود القصوى والدنيا للدرجات ووسيط مستوى القدرة للمجموعة (على مقياس بين ١ إلى ١٠٠). راجع أيكين ١٩٨٣ لمعلومات أكثر تفصيلاً عن هذا البرنامج.

٧- احتمالية قيام ممتحنين بأخطاء متماثلة: يستطيع هذا البرنامج اكتشاف الغش في الاختبارات الموضوعية حيث يقوم البرنامج بحساب الإحتمالية ثنائية الحد لقيام ممتحنين بإعطاء عدد محدد من الإجابات الخاطئة المتماثلة لمجموعة من فقرات أسئلة الخيارات المتعددة أو الصواب والخطأ. ويقوم المستخدم في حالة أسئلة الخيارات المتعددة بإدخال عدد الخيارات (البدائل) لكل فقرة وعدد الفقرات التي تمت إجابتها بشكل خاطئ من قبل كلا الممتحنين وعدد الفقرات التي اختار فيها الطالبان المموه نفسه. ويقوم المستخدم، في فقرات الصواب والخطأ (أو الفقرات ذات الخيارين)، بإدخال عدد الفقرات التي أخطأ فيها أحد أو كلا الممتحنين وعدد الفقرات التي أخطأ فيها كلا الممتحنين. راجع أيكين ١٩٩١ لمزيد من المعلومات التفصيلية حول النظرية التي يستند إليها هذا البرنامج.

برامج الفئة C تحليل الفقرات

١- مؤشرات الصعوبة والتمييز للفقرات ذات المرجعية المعيارية: يقوم هذا البرنامج القصير بحساب مؤشر صعوبة فقرات الاختبار (p) ومؤشر تمييز الفقرات (D) لعدد معين من الفقرات. يقوم المستخدم بإدخال العدد الكلي للفقرات وعدد الممتحنين في المجموعات العليا والدنيا (والتي يتم تحديدها

بناءً على المجموع الكلي للدرجات) الذين أجابوا على سؤال ما بشكل صحيح. يتم إدراج قيم P و D تبعاً لرقم الفقرة في النسخة المطبوعة.

٢- مؤشرات الصعوبة والتمييز للفقرات ذات المرجعية المحكية: يقوم هذا البرنامج القصير بحساب مؤشر صعوبة ومؤشر تمييز الفقرات للاختبارات ذات المرجعية المحكية. وتظهر قيم المؤشرين على شاشة العرض.

٣- تحليل الفقرات الثمانية وإحصائيات الدرجات الكلية واختبارات الدلالة: يقوم هذا البرنامج بحساب وتقييم ثمانية مؤشرات إحصائية لأداء n من الممتحنين في عدد من الفقرات الخيارات المتعددة m. تتضمن المؤشرات قياسات لما يلي:

- درجة صعوبة الفقرة.
- الأداء العام للممتحنين في الاختبار.
- التشابه بين إجابات الممتحن على فقرة ما و إجاباته على الفقرات الأخرى.
- التشابه بين إجابات الممتحن على فقرة ما و إجابات ممتحن آخر على نفس الفقرة.
- التشابه في اختيار المموهات.
- العلاقة بين الفقرة ومعدلات الدرجات الكلية (مؤشر تمييز الفقرة).
- الاختلافات بين معدلات الدرجات الكلية لفقرة ما.
- الاختلافات بين معدلات الدرجات الكلية للممتحنين.

يتم تعريف المؤشر واختبار الدلالة الموافق له، ويتم تخزين مصفوفة المعطيات (m x n المتحن × البند) في ملف معلومات منفصل.راجع أيكين (a١٩٨٩) لمزيد من المعلومات حول البرنامج.

٤- خصائص الفقرات ومنحنى إجابة الفقرات: يقوم هذا البرنامج برسم منحنيات خصائص الفقرات بناءً على المعطيات الأولية أو من خلال نموذج معين للإجابة على الفقرات. يقوم المستخدم بالإشارة إلى ما إذا كان يجب تحديد منحنى خصائص الفقرة من المعطيات الأولية أو من نموذج الإجابة على الفقرات ثم يقوم بإدخال العدد الكلي لمعدلات الدرجات وعدد المجالات ونقطة الوسط لكل مجال وعدد معدلات الدرجات في كل مجال وعدد الممتحنين الذين نجحوا في الفقرة في كل مجال. وفي حال اختيار أحد نماذج الإستجابة للفقرة، يتوجب على المستخدم أن يشير إلى ما إذا يرغب في استخدام نموذج راش أو النموذج ذو المؤشر الواحد أو نموذج ثنائي المؤشرات أو نموذج ثلاثي المؤشرات. يحدد بعد ذلك قيم المؤشرات ويظهر على الشاشة رسم لمنحنى خصائص الفقرة.

٥- نقطة الارتباط ثنائي التسلسل (Point biserial): يقوم هذا البرنامج بحساب معامل نقطة الارتباط ثنائية التسلسل لجملة من معدلات الدرجات المزاوجة (التي تم أخذها على شكل أزواج) على متغير ذو حدين (خيارين) أو اختيارات متعددة. يقوم المستخدم بإدخال العدد الكلي للممتحنين ووسطي درجات المحك والإنحراف المعياري لدرجات المحك والعدد الكلي للفقرات. كما يقوم المستخدم بالإشارة إلى العدد الكلي للممتحنين الذين نجحوا في فقرة ما ويدخل معدل درجاتهم على مقياس المحك. يتم تخزين معاملات النقطة ثنائية التسلسل تبعاً لترقيم الفقرات في ملف results كما يمكن طباعتها على شاشة العرض أو على الورق.

٦- تحليل فقرات الاختبار الاختيارية: يقوم هذا البرنامج بتقدير مؤشرات الصعوبة والتمييز وثبات مجموعة من فقرات الاختبار الاختيارية (تتألف هذه المجموعة من فقرتين أو أكثر) عندما يطلب من الممتحنين الإجابة على عدد محدد (أكثر من واحد) من الفقرات.

يقوم البرنامج بتحديد ما يلي:

- صعوبة الفقرات الاختيارية بناءً على عدد الممتحنين الذين أجابوا عليها
- الصعوبة النسبية لأزواج من الفقرات الاختيارية بناءً على المجموعات الفرعية من الممتحنين الذين أجابوا على الفقرتين معاً
- الارتباطات النقطية ثنائية التسلسل بين الفقرات الاختيارية التي تمت الإجابة عليها أو لم تتم الإجابة عليها ودرجات الممتحنين في الفقرات المطلوبة
- الارتباطات بين درجات الفقرات المطلوبة ودرجات الفقرات الاختيارية
- الارتباطات بين أزواج من الفقرات الاختيارية
- الحد الأدنى التقديري لثبات الفقرات الاختيارية عندما يكون ثبات الفقرات أو الأقسام المطلوبة أ- معلومة أو ب- غير معلومة
- تقديرات الحد الأدنى لثبات الفقرات الاختيارية بناءً على الارتباطات مع أزواج من الفقرات الاختيارية الأخرى عندما يكون ثبات الفقرات الاختيارية غير معلوم. تتألف المعطيات المدخلة من درجات الفقرات المطلوبة الإلزامية والفقرات الاختيارية لكل ممتحن مدرج في بند واحد من الملف المختار. ويتم التعبير عن الفقرات المحذوفة بأعداد أكبر من معدل درجات الأعلى المحتمل لكل فقرة. راجع أيكين (b١٩٨٩) للمزيد من التفاصيل عن الإجراءات الحسابية.

٧- تقييم عمل الفقرات التبايني: يقوم هذا البرنامج بحساب قيم دلتا من نسب الممتحنين من مجموعتين والذين نجحوا في فقرات محددة. كما يرسم البرنامج قيم دلتا للمجموعة الثانية بالنسبة لقيم المجموعة الأولى ويولد معادلة الإنحدار الموافقة لها (آنغوف وفورد، ١٩٧٣). كما يُجري البرنامج

اختبار مربعات – كاي لدلالة عمل الفقرات التفاضلي (DIF) باستخدام إجراءات شوينمان (١٩٧٩) وكاميللي (١٩٧٩) ومانتل وهاينتزل (١٩٥٩).

برامج الفئة D: المعايير واختيار العينات:

١- اختيار العينات الطبقي العشوائي: يقوم البرنامج باختيار عينة عشوائية متعددة الطبقات مكونة من m عنصراً من أصل تعداد عام للسكان مكون من n عنصراً. ويتناسب اختيار عدد العناصر من كل طبقة ومن كل بُعد تم تصنيف السكان على أساسه مع التعداد العام للسكان في كل طبقة. وفي حال وجود طبقة واحدة وبعد واحد، يتم اختيار عينة عشوائية بسيطة من مجموع السكان مكونة من عدد العناصر n. أما في حال تساوي عدد العناصر في العينة مع التعداد الكلي للسكان، فيتم إجراء إعادة ترتيب عشوائي لعدد العناصر n.

٢- الخطوط والرتب المئينية: يقوم هذا البرنامج بحساب التالي (وذلك إنطلاقاً من توزيع تكراري معلوم):

الرتب المئينية لمعدل درجات معطى أو الخط المئيني الموافق لرتبة مئينية معينة.

٣- الدرجات القياسية: يقوم هذا البرنامج، إنطلاقاً من توزيع تكراري معلوم، بحساب معدلات الدرجات القياسية z ومعدلات الدرجات المطبّعة znormalized ومعدلات الدرجات المشتقة T وZ المكافئة لمعدلات الدرجات تلك.

يقوم المستخدم بتحديد عدد مجالات الدرجات ونقطة الوسط للمجال الأول وعرض المجال، ثم يقوم المستخدم بتحديد تكرار الدرجات على كل من المجالات التي تتبع المجال الأول. وتتألف المخرجات التي تظهر على شاشة العرض من جدول يوضح نقطة الوسط والتكرار والرتبة المئينية المتوضعة في نقطة الوسط وقيم z وZ وz المطبّعة وقيم T لكل مجال.

٤- الاحتمالات والانحرافات الطبيعية: يمكن استخدام هذا البرنامج لحساب إما الإحتمالية الطبيعية لقيمة معلومة للدرجات z أو للدرجات z و T وCEEB والرتب من تسـعـة Stanine وNCE (مكافئ المنحنى الطبيـعي) الموافقةلاحتمالية تراكمية طبيعية معطاة.

برامج الفئة E: معاملات الثبات

١- معادلات سبيرمان - براون لحساب الثبات: يمكن استخدام هذا البرنامج لتقدير التالي:

- ثبات اختبار تم تطويله: يدخل المستخدم عدد الفقرات الاختبارية للاختبار الأصلي (قبل التطويل) وعدد الفقرات الكلي للاختبار بعد التطويل وثبات الاختبار الأصلي (غير المطوّل). تظهر قيمة ثبات الاختبار المطوّل على شاشة العرض.
- عدد الفقرات الاختبارية الإضافية اللازمة لرفع ثبات اختبار ما إلى قيمة محددة : يدخل المستخدم عدد الفقرات الاختبارية للاختبار الأصلي (قبل التطويل) وثبات الاختبار الأصلي (غير المطوّل) وقيمة الثبات المرغوب فيها للاختبار المطوّل. ويظهر عدد الفقرات التي يجب إضافتها لرفع قيمة ثبات الاختبار المطوّل إلى القيمة المرادة على شاشة العرض.

٢- معاملات ثبات كودر - ريتشاردسون: يقوم هذا البرنامج بحساب ثبات الإتساق الداخلي لاختبار ما من خلال معادلات كودر - ريتشاردسون رقم ٢٠ و٢١، ويقوم المستخدم بإدخال عدد الفقرات الاختبارية والوسط الحسابي لمجموع معدلات درجات الاختبار وتباين مجموع معدلات الدرجات ونسبة عدد الممتحنين الذين أجابوا على الفقرة الاختبارية بشكل صحيح (وذلك لكل فقرة اختبارية على حدة). تظهر معالات كودر- ريتشاردسون رقم ٢٠ و رقم ٢١ على شاشة العرض.

٣- المعامل ألفا: يقوم البرنامج بحساب المعامل ألفا من خلال المعطيات المدخلة عن طريق لوحة المفاتيح أو عن طريق ملف معطيات، والمعامل ألفا هو مقياس لثبات الإتساق الداخلي لاختبار تمنح فيه الدرجات على أساس النقاط المتعددة لكل فقرة (أي لا يعتمد مبدأ إعطاء نقطة واحدة للإجابة الصحيحة و٠ للإجابة الخاطئة). يقوم المستخدم بإدخال عدد الفقرات الاختبارية وعدد الممتحنين والدرجات التي حصل عليها كل فرد في كل فقرة. نظهر القيمة العددية للمعامل ألفا على شاشة العرض.

٤- معامل ثبات المصححين: يقوم هذا البرنامج بحساب معامل الارتباط بين الدرجات التي يضعها عدد من المصحيحين للورقة الاختبارية ذاتها، كما يحسب البرنامج الإرتباط بين مجموع الدرجات العام أو معدلات الدرجات التي توصل إليها أكثر من مصحح.

٥- معامل التوافق: يحدد هذا البرنامج درجة التوافق بين الرتب أو الدرجات التي يمنحها مصححان أو أكثر لورقة اختبارية ما. كما يحسب البرنامج بالإضافة إلى معامل التوافق وقيم مربعات كاي المقابلة والارتباط بين معدلي الرتب والارتباط المتوقع بين مجموعتين من الرتب.

٦- الخطأ المعياري للقياس : يقوم هذا البرنامج بحساب الخطأ المعياري للقياس ومجالات الثقة لـ ٩٠٪ و ٩٥٪ و٩٨٪ و ٩٩٪ لمعدل الدرجات الحقيقي لاختبار ما. ولتحديد الخطأ المعياري في القياس ومجال الثقة المقابل له، يقوم المستخدم بإدخال الانحراف المعياري للاختبار ومعامل الثبات.

٧- الثبات والخطأ المعياري لفرق معدلات درجات: يقوم هذا البرنامج بحساب الثبات والخطأ المعياري لفرق معدلات درجات. يدخل المستخدم الانحرافات المعيارية لاختبارين وثباتهما والارتباطات بينهما. عندها يظهر على شاشة

العرض الثبات والخطأ المعياري لفرق معدلات درجات. بعد ذلك، يسأل البرنامج المستخدم ما إذا كان يرغب بحساب مجال الثقة لفرق معدلات الدرجات. وفي حال الجواب بنعم، يطلب من المستخدم الدلالة على النسبة المئوية المطلوبة (٩٠، أو ٩٥، أو ٩٩) لدرجات الاختلاف وللفرق بين معدلي الدرجات. يظهر مجال الثقة لهذا الفرق على الشاشة ويُسأل المستخدم ما إذا كان يرغب بحل مسألة أخرى.

٨- الثبات والنسبة القصوى لثبات مركّب اختباري: يقوم البرنامج بحساب الثبات لمجموع أجزاء من اختبار مركب موزونه ويحدد الحد الأقصى للثبات والأوزان المقابلة له. يقوم المستخدم بإدخال عدد المتغيرات للأوزان، ومعامل الثبات والارتباط بين كل زوج من المتغيرات. ويمكن للبرنامج حساب إما : (١) معامل الثبات والخطأ المعياري للقياس المرتبط بأي مجموعة من الأوزان المرجحة للمتغيرات يدخلها المستخدم في البرنامج. و(٢) الحد الأقصى لثبات المركب الاختباري والخطأ المعياري للقياس والأوزان لكل متغير. راجع أيكين (١٩٨٨) للمزيد من المعلومات حول البرنامج.

برامج الفئة F: الاختبارات التحصيلية

١- الاختبارات الحسابية الموضوعة بشكل عشوائي: يقوم هذا البرنامج ببناء وإجراء ووضع درجات اختبار واحد أو أكثر يتألف من عشر فقرات اختبارية وتم وضعه عشوائياً في الجمع أو الطرح أو الضرب أو القسمة أو للرفع على أس للأعداد المؤلفة من خانة واحدة أو خانتين أو خانة واحدة وخانتين في آن واحد.

٢- اختبار التهجئة واستعمال الكلمات: يقوم هذا البرنامج بإجراء ووضع درجات اختبار مؤلف من ثلاثة أجزاء:

- الجزء الأول هو اختبار لمدى المعرفة للتهجئة الصحيحة لـ ٢٨ كلمة يشيع الخطأ في تهجئتها في الكتابة الرسمية؛
- الجزء الثاني هو اختبار لمدى معرفة التهجئة الصحيحة لمجموعة من ٨ كلمات؛
- الجزء الثالث هو اختبار على ٢٠ زوج من الكلمات ذات معانٍ مختلفة.

يعطي البرنامج الدرجات الأولية والنسبة المئوية للإجابات الصحيحة لكل جزء من الاختبار وللاختبار بكامله.

٣- حقائق عن التقدم في السن والحياة الآخرة: يقدم البرنامج اختباراً مكوناً من ٥٠ فقرة من نوع الصواب أو الخطأ حول التقدم في السن والحياة الآخرة. تطبع نسبة الإجابات الصحيحة والدرجات المحتملة المقابلة لها على شاشة العرض.

٤- حقائق عن الموت والاحتضار: يقوم هذا البرنامج بإجراء وإعطاء معدلات الدرجات لاختبار مسّاح مؤلف من ٥٠ فقرة صواب أو خطأ حول الموت والاحتضار. تطبع نسبة الإجابات الصحيحة والدرجات المحتملة المقابلة لها على شاشة العرض.

٥- اختبار حول المصطلحات النفسية: يقوم هذا البرامج باختيار وبشكل شبه شامل عبر ٥٥ فقرة اختبارية ثنائية الخيار معرفة المصطلحات المستخدمة في الدورات التمهيدية في علم النفس. يسمح البرنامج للممتحن بأن يقوم بمراجعة وتغيير إجاباته على الأسئلة واستعراض الإجابات الصحيحة في نهاية الاختبار. يظهرعلى شاشة العرض معدل الدرجات الأولية (الرقم الصحيح) ومعدل الدرجات المصحح من أثر التخمين والنسب المئوية لمعدلات الدرجات الأولية ومعدلات الدرجات المصححة من أثر التخمينمن من معدلات الدرجات المحتملة للاختبار.

٦- اختبار سرعة القراءة ودرجة الاستيعاب للنصوص: لقد تم إعداد هذا الاختبار للتعرّف على السرعة ومدى الاستيعاب عند قراءة نص قصير. يطلب من المتحن قراءة نص ما تبعاً لسرعته في القراءة وضغط أي مفتاح عند الانتهاء. يتم اختبار المتحن بأربعة أسئلة من نوع الخيارات المتعددة، ويطلب من المتحن إدخال الرقم المقابل للإجابة الصحيحة. يظهر زمن القراءة (بالثواني) ومعدل الدرجات الأولية ومعدل الدرجات المئوي على شاشة العرض (تم أخذ النص والفقرات الامتحانية من تمارين اختبار GED الرسمي لعام ١٩٩٣ ص٣٧، وأعيدت طباعتها بموافقة مجلس التعليم الأمريكي).

٧- اختبار سرعة ودقة الطباعة: هذا اختبار لسرعة ودقة طباعة نص قصير. هذا يظهر نص في أعلى شاشة العرض ويطلب من المتحن طباعته في أسفل الشاشة والضغط على مفتاح الإدخال عند الانتهاء. يظهر على شاشة العرض الزمن (بالثواني) وعدد الأخطاء والنسبة المئوية للإجابات الصحيحة. (تم أخذ النص من تمارين كتيب الوصف العام لاختبار GED 1996-1997 ص٢٣، أعيدت الطباعة بموافقة مجلس التعليم الأمريكي).

٨- الاختبارت التحصيلية القياسية: يقدم هذا الاختبار البرنامج معلومات تعريفية عامة بالاختبارات التحصيلية المنشورة مبوبة في عشر فئات:

- اختبارات المسح للقراءة.
- اختبارات التشخيص والجاهزية للقراءة.
- اختبارات اللغة الإنجليزية.
- اختبارات اللغات الأجنبية.
- اختبارات الرياضيات.
- اختبارات العلوم.
- اختبارات التاريخ والدراسات الاجتماعية.

● اختبارات مهارات الفن والموسيقى والمكتبات.

● اختبارات الأعمال والاختبارات المهنية.

برامج الفئة G: اختبارات القدرات الحركية والإدراكية - الحسية:

١- اختبار المدى البصري: يقوم هذا البرنامج بقياس درجة الحساسية لأجزاء مختلفة من المجال البصري، وهو اختبار طبي للعيون شائع الاستخدام. يطلب من الممتحن ضبط شدة السطوع على شاشة العرض حتى تصبح رؤية الكلمات أمراً صعباً للغاية ثم يطلب من الممتحن التركيز على بقعة مضيئة صغيرة في وسط الشاشة. ومع كل محاولة من المحاولات الـ١٠٨ تظهر بقعة أخرى مضيئة في مكان ما من الشاشة وعلى الممتحن أن يضغط على مفتاح المسافة حالما يتمكن من رؤية البقعة المضيئة الثانية. تظهر المحفّزات البصرية بشكل عشوائي في ١٢ قطاعاً مختلفاً من الشاشة، ويتم تسجيل إحداثيات المحفِّز البصري وبُعده عن مركز الشاشة ورقم القطاع الذي ظهر فيه والوقت الذي استغرقه الممتحن في الإجابة وذلك لكل محاولة على حدة. وعند الانتهاء من الاختبار يتم حساب متوسط زمن الاستجابة لكل قطاع. ويمكن لتحليل ثنائي الاتجاه (٣-٤) للتباين أن يوضح كيفية تغير زمن الإجابة بتغير موقع المحفز في المجال البصري.

٢- اختبار تمييز اللون: هذا اختبار للقدرة على التمييز بين ألوان أزواج من الوجوه. يطلب من الممتحن الإشارة إلى ما إذا كانت ألوان أزواج الوجوه متماثلة (s) أو مختلفة (d) ويتم تقديم ٢٢ زوجاً من الوجوه (١١) منها ذات ألوان متماثلة و(١١) ذات ألوان مختلفة. تطبع على الشاشة معدلات درجات المجيب بالنسبة لفئات: "أحمر- أخضر" و"أزرق - أصفر" و"الأخطاء المتطابقة" إضافة إلى النسبة المئوية من معدل الدرجات الكلي المحتمل والمتمثل بكل واحد من معدلات الدرجات للفئات الثلاثة المذكورة.

٣- اختبار حدة السمع: يقوم هذا البرنامج بتحديد الحد الأعلى وأعلى تردد للحساسية لشدة الصوت لدى الممتحن. يسمع الممتحن سلاسل مختلفة من النغمات تزداد شدة الصوت فيها بشكل تصاعدي، ويطلب من الممتحن في السلسلة الأولى من النغمات الضغط على مفتاح المسافة عندما يتوقف عن سماع النغمات. أما في السلسلة الثانية فيطلب من الممتحن الضغط على مفتاح المسافة عندما يصل علو النغمة إلى ما يجده حداً أقصى. ومن الجدير بالذكر أن عمر الممتحن وحالته الصحية وغيرها من العوامل تلعب دوراً في الحد المطلق لسماع وتردد الحساسية القصوى لشدة الصوت (يتم حساب التردد بوحدة الهيرتز).

٤- اختبار تمييز الشدّة: يتم تقديم ١٤ زوجاً من النغمات كل زوج يتكون من نغمتين لهما إما نفس التردد أو لكل منهما تردد مختلف. ويطلب من الممتحن الضغط على "١" إذا كانت النغمة الأولى أعلى من حيث الشدة والضغط على "٢" إذا كانت الثانية أعلى. تطبع على شاشة العرض عدد الأزواج التي أجاب عليها الممتحن بشكل صحيح والنسبة المئوية لمعدلات الدرجات من أعلى معدل درجات ممكن.

٥- اختبار البحث عن الأرقام: يمثل هذا الاختبار مقياساً لسرعة الملاحظة و الدقة في تحديد المكان والاستجابة لمجموعة من الأرقام. يقوم المُمتحَن بالبحث في مصفوفة أرقام (١٠ × ١٠) مولدة عشوائياً من الأرقام من (٠) وحتى (٩٩) لتحديد مكان رقم معين وبعد ذلك يقوم بإدخال رقم السطر والعمود اللذين يقع فيهما هذا الرقم. يمكن للقائم على الاختبار (أو الممتحَن) أن يقوم بتحديد الحد الأقصى للرقم الواجب البحث عنه والواقع في المجال بين (٠) و(٩٩). وتنص تعليمات الاختبار على ما يلي: انتظر حتى تظهر مصفوفة الأرقام وحتى يتم تعيين الرقم المطلوب ثم أدخل رقم السطر والعمود اللذين يقع فيهما الرقم المطلوب.

٦- اختبار زمن ردّات الفعل دوندرز (أ) و(ب) و(ج): يقوم هذا البرنامج بقياس زمن ردة الفعل البسيطة و المنفردة باستخدام إجراءات دوندرز من الفئات (أ) و(ب) و(ج). وفي الفئة (أ) يقوم المتحَن بالضغط على مفتاح محدد بأسرع وقت ممكن بعد ظهور حرف تحفيزي على شاشة العرض. أما في الفئة (ب) فيظهر واحد من أصل حرفين تحفيزيين على شاشة العرض و على المتحَن الضغط على المفتاح المحدد عند ظهور أحدهما فقط.

أما في الفئة (ج) فيقوم المتحن بالضغط على مفتاح محدد عند ظهور أحد حرفين تحفيزيين على شاشة العرض وعلى مفتاح آخر عند ظهور الحرف الآخر تظهر بعد ذلك على شاشة العرض وفي ملف النتائج أزمنة ردّة الفعل للفئات (أ) (الزمن الأساسي)و(ب) و(ج) إضافة لزمن التعرف وزمن الاختيار الذي تم حسابه.

٧- اختبار السرعة والصحة للعاملين في الأعمال الكتابية: يتم في هذا الاختبار تقديم ٣٠ زوجاً من الأسماء أو الأرقام بمعدل زوج كل ثانيتين، ويطلب من المتحَن الضغط على المفتاح "s" إذا كان الاسمان أو الرقمان متشابهين وعلى المفتاح "d" إذا كان الاسمان أو الرقمان مختلفين. يتم إعلام المتحَن بأنه لديه دقيقة واحدة للإجابة على الفقرات الثلاثين لذا يجب الإجابة على الفقرات بأسرع ما يمكن. وبعد أن يتم تقديم كافة الفقرات يظهر على الشاشة معدل الدرجات النهائي (الصح و الخطأ) والنسب المئوية المقابلة لها.

برامج الفئة H: اختبارات القدرات الإدراكية

١- الأحاجي (Anagrams): صُمم هذا البرنامج لتحديد الزمن اللازم لتركيب كلمة ذات معنى من كل مجموعة من المجموعات الإثنتي عشر المكونة من خمس أحرف مبعثرة وذلك مع أو بدون أدلة تمهيدية. وفي حالة إعطاء دليل (clue)، يتم إخبار المتحَن بأن الكلمة إما لون أو لها علاقة بالطعام أو العمل،

أمـا في حـالة عـدم إعطاء دليل (no-clue) فـلا يتم إعطاء أي دليل. تخـزن الأحـاجي فـي الملف (scrabble). يتـم تخـزين عـدد المحـاولات وزمن الإجـابة بالثواني وإجابات الممتحَن في الملف results حيث تظهر على الشاشة و يمكن أيضاً طباعتها على الطابعة.

٢- اختـبـار المدى الرقـمي: يقـوم هذا البـرنامج باخـتبـار القـدرة على اسـتدعـاء سلسلة من الأرقام المفردة المولدة عشـوائيـاً والمقـدمة بمعـدل رقم واحـد في الثانية والقدرة على تكرارها بنفس الترتيب الذي تم عرضها فيه (digits for-ward) أو بالترتيب العكسي (digits backward). تنص تعليمات الاختبار على ما يلي وذلك في حـالة (digits forward): سـوف تظهـر سلسلة مـن الأرقـام المفردة من ٢ وحتى ٩ في منتصف الشاشة كل رقم على حدة. حاول أن تحفظ الأرقام مـن كل سلسلة حتى تتمكن من إعادة طباعتها بنفس التـرتيب الذي تم تقـديمها فيـه. وتنص تعليـمات الاخـتبـار في حـالة (digits backward) على التـالي: سـوف تظهـر سلسلة من الأرقـام المفردة مـن ٢ و حـتى ٩ في منتـصف الشاشة كل رقم على حدة، حاول أن تحـفظ الأرقام من كل سلسلة حتى تتمكن من إعادة طبـاعتها بالتـرتيب المعاكس للترتيب الذي تم تقديمهـا فيـه. وفي كلا الحالتين، يتم إعلام الممتحَن بالتعليمات التالية:

بعد أن يتم تقديم كل سلسلة من الأرقام تظهر على الشاشة ومباشرة الرسالة: "Digit?"، عنـدهـا قم بطبـاعـة الـرقم التـالي ثم قم بالضـغط على مـفـتـاح الإدخـال. وإذا قـام المـمـتـحَن بخطأ في أحـد السـلاسل فـيـتم تقـديم سلسلة جديدة بنفس الطول. يمنح الممتحَن درجة واحدة لكل إجابة صحيحة وبعد أن يتم تقـديم السـلاسل الثـمـانيـة لكل حـالة يظهـرعلى شـاشـة العـرض مـعـدل الدرجات لتلك الحالة.

٣- سـلاسل الأعـداد: يقـوم هـذا البـرنامج بتـقـديم عـشـر سـلاسل مكونة من ست

أرقام و يتوجب على الممتحَن تحديد الرقم التالي من السلسلة عبر تطبيق قاعدة حسابية. يتم إعطاء التعليمات التالية:

هذا اختبار لقدرتك على استنتاج الأنماط لمجموعة من الأرقام. سوف يكون هناك عشر محاولات بالمجمل. في كل محاولة سوف تظهر في وسط الشاشة سلسلة من ست أرقام. إن مهمتك ستكون تحديد ماذا يجب أن يكون الرقم التالي (السابع) في السلسلة ثم طباعته والضغط على مفتاح الإدخال. الوقت المتاح هو ٣٠ ثانية لكل سلسلة من السلاسل العشرة. يتم إعلام الممتحَن بعد كل محاولة ما إذا كانت إجابته صحيحة أم خاطئة ويتم تخزين رقم المحاولة وزمن الإجابة وكون الإجابة صحيحة أم خاطئة. يمكن طباعة هذه المعلومات بالإضافة إلى معدل الدرجات الكلي للممتحَن على الشاشة أو على الورق.

٤- مسائل في البصيرة: يختبر هذا البرنامج القدرة على حل مسائل مكونة من مجموعة من ١٢ كلمة. يتم تخزين هذه المسائل في الملف "probs" والحلول العددية في مفتاح الملف. يطلب من الممتحَن قراءة كل مسألة وطباعة و إدخال الإجابة العددية لها(١ أو ٢ أو ٣ إلخ). وفي نهاية الاختبار يظهر على الشاشة معدل الدرجات والنسبة المئوية الصحيحة.

٥- اختباران للقدرات الإدراكية: الاختباران في هذا البرنامج هما مقاييس خاصة للغرض للقدرات الإدراكية والتي لا توجد لها معايير ثابتة. يتكون الاختبار الأول "اختبار المتشابهات اللغوية" من ٢٠ فقرة اختبارية ذات خمس خيارات مصممة لقياس جانب واحد من الذكاء - تعلم العلاقات والارتباطات. أما الاختبار الثاني "اختبار الكلمات المترابطة" فهو اختبار إتمام ٢٠ فقرة اختبارية تم تصميمها تبعاً لنمط اختبار المترابطات البعيدة يقيم القدرات الإبداعيةوهو في الواقع اختبار للقدرة اللغوية أو للسهولة في التعامل مع الكلمات. يظهر معدل درجات الممتحَن (الرقم الكلي الصحيح) في كل اختبار على شاشة العرض عند انتهاء الاختبار.

٦- مسائل برج هانوي: هذا البرنامج مستوحى من مسألة برج هانوي التقليدية إلا أنه يستخدم الأرقام بدلاً من المكعبات. يطلب من الممتحَن تحريك كافة الأعداد في العمود (أ) إلى العمود (ب) على شرط عدم وضع رقم ما فوق آخر أصغر منه وذلك باستخدام عمود (ج) كمرحلة وسيطة أو كمخزن أثناء عملية نقل الأعداد من العمود (أ) إلى العمود (ب). تتم عملية نقل الأرقام من خلال إملاء الفراغات في الجملة غير المكتملة: "انقل الرقم العلوي في العمود ـــــ إلى العمود ـــــ. إطبع الأحرف أ، ب، ت،في كل فراغ. إضغط على مفتاح الإدخال بعد القيام بطباعة الحرف. وغالباً ما يبدأ الاختبار بمسألة بسيطة كمسألة مكونة من ثلاثة أرقام و يتطور إلى أربعة أو خمسة أو ستة أرقام ونظرياً حتى عشرة أرقام. إن عدد الحركات في الحل الأمثل هوحيث n هو عدد الأرقام (المكعبات) الواجب تحريكها. ويشار إلى الأخطاء في الحركات عن طريق صوت رنين لتنبيه الممتحن. وبعد التوصل إلى الحل الصحيح يظهر على الشاشة عدد الحركات التي قام بها الممتحَن والعدد الأمثل للحركات (العدد الأدنى).

٧- فهرس اختبارات القدرات الإدراكية: يقدم هذا البرنامج معلومات توصيفية عن المئات من اختبارات القدرات العامة والخاصة مقسمة إلى عشر فئات:

- اختبارات الذكاء الفردية
- اختبارات الذكاء الجماعية
- اختبارات القبول المهنية والجامعية
- اختبارات الارتقاء
- اختبارات ذوي الاحتياجات الخاصة
- اختبارات الذاكرة
- الاختبارات العصبية - النفسية.
- اختبارات الاستعدادات المتعددة.
- اختبارات الاستعدادات الخاصة.
- اختبارات المواهب.

الملحق ب

دور نشر وتوزيع أدوات التقدير النفسية والتعليمية

تحتوي القائمة أدناه على أسماء وعناوين دور النشر والتوزيع الكبيرة ولكن لا يمكن الجزم بكون هذه القائمة كاملة. يمكن الحصول على أسماء وعناوين أخرى في الصفحات ١١٤٥-١١٥٨ من كتاب الاختبارات المنشورة ٤.

Academic Therapy Publications
20 Commercial Boulevard
Novata, CA 94949-6191

ACS DivCHED Examination Institute
Oklahoma State University
107 Physical Sciences
Stillwater, OK 74078

The Alemany Press
P.O. Box 7604
West Trenton, NJ 08628

American College Testing Program
(ACT)
P.O. Box 168
Iowa City, IA 52243

American Council on Education
One Dupont Circle NW, Suite 250
Washington, DC 20036-1163
202-939-9490

American Guidance Service (AGS)
4201 Woodland Road
P.O. Box 99
Circle Pines, MN 55014-1796
800-328-2560

American Orthopsychiatric
Association, Inc.
49 Sheridan Avenue
Albany, NY 12210

Assessment Systems Corporation
2233 University Avenue, Suite 200
St. Paul, MN 55114-1629
612-647-9220

The Australian Council for
Educational Research
19 Prospect Hill Road
Camberwell Vic 3124, Australia
(03) 9277 5656

Ballard & Tighe
480 Atlas Street
Brea, CA 92621

Behavior Data Systems, Ltd.
P.O. Box 44256
Phoenix, AZ 85064-4256
800-231-2401

Behavior Science Systems, Inc.
P.O. Box 580274
Minneapolis, MN 55448
612-929-6220

Center for Applied Linguistics
1118 22nd Street NW
Washington, DC 20037

Consulting Psychologists Press, Inc. (CPP)
3803 East Bayshore Road
P.O. Box 10096
Palo Alto, CA 94303
800-624-1765

CPPC
4 Conant Square
Brandon, VT 05733
800-433-8234

C.P.S., Inc.
P.O. Box 83
Larchmont, NY 10538
800-433-8324

CTB/McGraw-Hill
20 Ryan Ranch Road
Monterey, CA 93940
800-538-9547

Curriculum Associates, Inc.
5 Esquire Road
North Billerica, MA 01862-2589

Denver Developmental Materials, Inc.
P.O. Box 9619
Denver, CO 80206-9011

DLM Resources
One DLM Park
Allen, TX 75002
800-527-4747

Educational and Industrial Testing Service (EdITS)
P.O. Box 7234
San Diego, CA 92167
800-416-1666

Educational Developmental Corporation
P.O. Box 470663
Tulsa, OK 74147-0663

Educational Evaluation Enterprises
AWRE
Newnham, Gloucestershire GL14 1ET
England

Educational Testing Service (ETS)
ETS Test Collection (30-B)
Rosedale Road
Princeton, NJ 08540

Educators Publishing Service, Inc.
31 Smith Place
Cambridge, MA 02138-1100

GED Testing Service
One Dupont Circle NW
Washington, DC 20036

Grune & Stratton, Inc.
465 South Lincoln Drive
Troy, MO 63379

Hawthorne Educational Services, Inc.
800 Gray Oak Drive
Columbia, MO 65201
800-542-1673

Hilson Research, Inc.
P.O. Box 150239
Kew Gardens, NY 11415-0239
800-926-2258

Hodder & Stoughton Education
Headline PLC, 338 Euston Road
London NW1 3BH
England

Human Sciences Research Council
P.O. Box 32410, 1
Braamfontein 2017
South Africa

Industrial Psychology International, Ltd.
4106 Firestone Road
Champaign, IL 61821
217-398-1437

Institute for Personality and Ability Testing (IPAT)
P.O. Box 1188
Champaign, IL 61824-1188
800-225-IPAT

Psychologists and Educators, Inc.
P.O. Box 513
Chesterfield, MO 63006

Psychometric Affiliates
Box 807
Murfreesboro, TN 37133-0807
615-898-2565

Publishers Test Service, CTB/McGraw-Hill
20 Ryan Ranch Road
Monterey, CA 93940
800-538-9547

Ramsey Corporation, Boyce Station Offices
1050 Boyce Road
Pittsburgh, PA 15241

Reitan Neuropsychology Laboratories
2920 South 4th Avenue
Tucson, AZ 85713-4819
602-882-2022

Richardson, Bellows, Henry & Co., Inc.
1140 Connecticut Avenue NW
Washington, DC 20036

Riverside Publishing
425 Spring Lake Drive
Itasca, IL 60143-2079
800-323-9540

Scholastic Testing Service, Inc. (STS)
480 Meyer Road
P.O. Box 1056
Bensenville, IL 60106-1617
800-642-6STS

Science Research Associates (SRA)
155 North Wacker Drive
Chicago, IL 60606

Selby MillSmith Ltd.
30 Circus Mews
Bath BA1 2PW
United Kingdom
01225 446655

Sigma Assessment Systems, Inc.
P.O. Box 610984
Port Huron, MI 48061-0984
800-265-1285

Slosson Educational Publications, Inc.
P.O. Box 280
East Aurora, NY 14052-0280
800-828-4800

SOI Systems
P.O. Box D
45755 Goodpasture Road
Vida, OR 97488
541-896-3936

Special Child Publications
P.O. Box 33548
Seattle, WA 98133

SRA/London House
9701 Higgins Road
Rosemont, IL 60018
800-221-8378

Stoelting
620 Wheat Lane
Wood Dale, IL 60191
630-860-9700

U.S. Employment Service, Division of Program Planning and Operations
U.S. Department of Labor
601 D Street NW
Washington, DC 20213

U.S. Military Entrance Processing Command Testing Directorate
2500 Green Bay Road
North Chicago, IL 60064

Western Psychological Services (WPS)
12031 Wilshire Boulevard
Los Angeles, CA 90025-1251
800-648-8857

Wide Range, Inc.
P.O. Box 3410
Wilmington, DE 19804-0250
800-221-WRAT

Wonderlic Personnel Test, Inc.
1509 North Milwaukee Avenue
Libertyville, IL 60048-1380
800-963-7542

Jastak Associates, Inc.
P.O. Box 3410
Wilmington, DE 19804-0250
800-221-WRAT

Jossey-Bass
615 Montgomery Street
San Francisco, CA 94111

Lafayette Instrument Company
P.O. Box 5729
3700 Sagamore Parkway North
Lafayette, IN 47903

Law School Admission Council
11 Dupont Circle NW, Suite 150
Washington, DC 20036

MacMillan Education
Houndsmills
Basingstoke, Hampshire RG21 2XS
England

Martin M. Bruce
50 Larchwood Road
Box 248
Larchmont, NY 10538
914-834-1555

McGraw-Hill/London House
9701 West Higgins Road
Rosemont, IL 60018-4720
800-237-7685

Merrill (Charles E.) Publishing Company
1300 Alum Creek Drive
Box 508
Columbus, OH 43216
800-848-1567

Metritech, Inc.
4106 Fieldstone Road
P.O. Box 6489
Champaign, IL 61826-6479
217-398-4868

MHS
908 Niagara Falls Boulevard
North Tonawanda, NY 14120-2060

Moreno Educational Co.
7050 Belle Glade Lane
San Diego, CA 92119

National Council on Economic Education
1140 Avenue of the Americas
New York, NY 10036

NCS Assessments
P.O. Box 1416
Minneapolis, MN 55440
800-627-7271

Nelson Canada
1120 Birchmount Road
Scarborough Ontario M1K 5G4
Canada

New Zealand Council for Educational Research
Education House West
178-182 Willis Street
Box 3237
Wellington, New Zealand

NFER-Nelson Publishing Co., Ltd.
Darville House
2 Oxford Road East
Windsor, Berkshire SL4 1DF
England

Personnel Press
191 Spring Street
Lexington, MA 02173

PRO.ED
8700 Shoal Creek Boulevard
Austin, TX 78757-6897
512-451-3246

Psychological and Educational Publications, Inc.
1477 Rollins Road
Burlingame, CA 94010-2316
800-523-5775

Psychological Assessment Resources, Inc. (PAR)
P.O. Box 998
Odessa, FL 33556
800-331-TEST

Psychological Corporation (The)
555 Academic Court
San Antonio, TX 78204-2498
800-211-8378

Psychological Publications, Inc.
P.O. Box 3577
Thousand Oaks, CA 91359-0577
800-345-TEST

Psychological Test Specialists
Box 9229
Missoula, MT 59807

الملحق ج

الموزعون التجاريون لمنتجات الحاسب ولخدمات التقدير النفسي والتعليمي

American College Testing
2201 North Dodge Street
Iowa City, IA 52243
319-337-1000

American Guidance Service
4201 Woodland Road
P.O. Box 99
Circle Pines, MN 55014-1796
800-328-2560

Assessment Systems Corporation
2233 University Avenue, Suite 200
St. Paul, MN 55114-1629
612-647-9220

Behavior Data Systems, Ltd.
P.O. Box 32938
Phoenix, AZ 85064-2938
800-231-2401

Caldwell Report
1545 Sawtelle Boulevard, Suite 14
Los Angeles, CA 90025
310-478-3133

Consulting Psychologists Press, Inc.
3803 East Bayshore Road
P.O. Box 10096
Palo Alto, CA 94303
800-624-1765

Educational and Industrial Testing Service
P.O. Box 7234
San Diego, CA 92167
619-222-1666

Hilson Research, Inc.
P.O. Box 150239
Kew Gardens, NY 11415-0239
800-926-2258

Institute for Personality & Ability Testing, Inc.
P.O. Box 1188
Champaign, IL 61824-1188
800-225-4728

Integrated Professional Systems, Inc.
5211 Mahoning Avenue, Suite 135
Youngstown, OH 44515
216-799-3282

Century Diagnostics, Inc.
2101 East Broadway, Suite 22
Tempe, AZ 85282
602-966-6006

CFKR Career Materials, Inc.
11860 Kemper Road, Unit 7
Auburn, CA 95603
916-880-2357

Clinical Psychometric Research
P.O. Box 619
Riderwood, MD 21139
301-321-6165

Computerized Psychological Diagnostics, Inc.
1101 Dove Street, Suite 225
Newport Beach, CA 92660
714-833-7931

Conover Company Ltd.
1050 Witzel Avenue
Oshkosh, WI 54901
414-231-4667

NFER-Nelson Publishing Company Ltd.
2 Oxford Road East
Windsor, Berkshire, SL4 1DF
England

Planet Press
P.O.Box 3477
Newport Beach, CA 92663-3418
714-650-5135

PRO.ED
8700 Shoal Creek Boulevard
Austin, TX 78757-6897
512-451-3246

Psychological Assessment Resources, Inc.
P.O. Box 998
Odessa, FL 33556
800-331-TEST

Psychological Corporation (The)
555 Academic Court
San Antonio, TX 78204-2498
800-228-0752

Psychological Psoftware Company
12486 Brickellia
San Diego, CA 92129
619-484-8877

Jastak Associates, Inc.
P.O. Box 3410
Wilmington, DE 19804-0250
800-221-WRAT

Life Science Associates
One Fenimore Road
Bayport, NY 11705
516-472-2111

Metritech, Inc.
4106 Fieldstone Road
P.O. Box 6489
Champaign, IL 61826-6479
800-747-4868

Multi-Health Systems, Inc.
908 Niagara Falls Boulevard
North Tonawanda, NY 14120-2060
800-456-3003

NCS Assessments
5605 Green Circle Drive
P.O. Box 1416
Minneapolis, MN 55443
800-NCS-7271

Slosson Educational Publications, Inc.
P.O. Box 280
East Aurora, NY 14052-0280
800-828-4800

StatSoft
2300 East 14th Street
Tulsa, OK 74104-4442
918-583-4149

Test Agency
Cournswood House, North Dean
High Wycombe, Bucks HP14 4NW
England
44-24-024-3384

University Associates
8517 Production Avenue
San Diego, CA 92121-2280
619-578-5900

Psychologistics, Inc.
P.O. Box 033896
Indialantic, FL 32903
407-259-7811

Risk & Needs Assessment, Inc.
P.O. Box 32818
Phoenix, AZ 85064-2818
800-231-2401

Riverside Publishing Company (The)
8420 Bryn Mawr Avenue
Chicago, IL 60631
800-767-TEST

Selby MillSmith
30 Circus Mews
Bath BA1 2PJ
Avon, England
(225) 446655

Sigma Assessment Systems, Inc.
P.O. Box 610984
Port Huron, MI 48061-0984
800-265-1285

Vocational Research Institute
1528 Walnut Street, Suite 1502
Philadelphia, PA 19102
800-874-5387

Western Psychological Services
12031 Wilshire Boulevard
Los Angeles, CA 90025-1251
800-648-8857

Wonderlic Personnel Test, Inc.
1509 North Milwaukee Avenue
Libertyville, IL 60048-1380
800-963-7542

قائمة المصطلحات والمفردات

اختبار القدرة: هو اختبار يقيس درجة قدرة الفرد على أداء مهام معينة أو احتراف مهنة معينة.

الاستعداد الأكاديمي: ويطلق عليه أيضاً اسم الاستعداد المدرسي، وهو القدرة على التعامل بنجاح مع المهام المدرسية أو ما يشابهها. تختبر معظم اختبارات الذكاء الاستعداد الأكاديمي عند الأفراد.

التحصيل: وهو درجة نجاح أو إنجاز فردٍ ما في مجال أو مشروع معين؛ وكذلك الدرجة المعطاة لاختبار للتحصيل.

السلوك التكيفي: وهو درجة قدرة الفرد على التواصل الفعال والمناسب مع البيئة المحيطة به.

الاختبار التكيفي: وهو إجراء اختباري غالباً ما يتم عن طريق الحاسب وتتنوع فيه الأسئلة والمهام تبعاً للقدرة المتوقعة للممتحن أو تبعاً لصفات الممتحن الخاصة أولإجاباته على الأسئلة السابقة.

التأثير المضاد (السلبي): وهو كل ما له تأثير سلبي على قبول بعض الأفراد المنتمين إلى جماعات معينة في بعض الوظائف والمؤسسات التعليمية.

المقاييس العاطفية: وهي اختبارات أو أدوات اختبار مصممة لقياس المزاج والشاعرية والاهتمامات والمواقف وأنماط الشخصية وغيرها من أنواع السلوك العاطفي للأفراد.

الدرجات المكافئة للعمر: راجع المعيار العمري.

المعيار العمري: وهو وسيط درجات اختبار تحصيلي أو اختبار للاستعداد أُجري لأطفال من عمر معين.

ثبات الصورة المكافئة: وهو مؤشر للثبات (معامل التكافؤ) يُحدد عن طريق الربط

بين درجات الفرد التي حصل عليها في صورة معينة للاختبارمع الدرجات التي حصل عليها في صورة أخرى من الاختبار ذاته.

التقييم البديل: وهي طريقة غير تقليدية للتقييم تعتمد على الطلب من المتحن أن ينتج أو يولد استجابات إبداعية لفقرات الاختبار بدلاً من الإجابة على فقرات محدودة الإجابات.

اختبار المتشابهات: وهو اختبار يتطلب من المتحنين تحديد العلاقات إما من حيث التشابه أو الإختلاف بين شيئين أو أكثر. مثال: الأحمر بالنسبة للورود هو كماذا بالنسبة للبنفسج: ١- الأزرق ٢- الأخضر ٣- البرتقالي ٤- الأصفر.

طريقة وضع الدرجات التحليلية: وهي إجراء لوضع الدرجات في الاختبارات المقالية يتم فيه تقييم إجابات المتحنين بالنظر إلى عاملي الأسلوب والمحتوى.

اختبار الفقرات المثبتة (anchor test): وهي عدد من فقرات الاختبار تبقى ثابتة في صور متعددة من ذات الاختبار وتساعد في معادلة درجات كافة صور الاختبار.

الاستعداد: وهو القدرة على تعلم كيفية تأدية مهمة ما أو مهارة معينة. وقد كان المعتقد السائد في الماضي هو أن الاستعداد قدرة تولد مع الفرد ولا تُكتسب بالخبرة والتدريب.

اختبار الاستعداد: وهو مقياس للقدرة على الإستفادة من تدريب أو خبرة إضافية تمكن الشخص من أن يصبح أكثر براعة أو خبرة. وتفيد الفقرات المختلفة في اختبارات الاستعداد في التنبؤ بالأداء المستقبلي للأفراد في مجالات العمل والتعليم.

التقدير: وهو تقييم وجود أو حجم صفات شخصية معينة. ويتضمن تقدير assessment السلوك البشري والعمليات العقلية عدداً من الإجراءات مثل الملاحظة والمقابلات وسلالم التقدير وقوائم الرصد ومقاييس السمات وتقنيات التنبؤ والاختبارات.

مقياس الصوت: وهو أداة لقياس الحدة السمعية يقوم على تقديم أنغام صافية تتعدد درجات عمقها وذبذبتها ضمن المجال السمعي العادي. يتم اختبار السمع في الأذنين، ثم تستخدم النتائج في رسم مخطط بياني سمعي يحدد حدة المتحن السمعية بالنسبة لكل ذبذبة وكل أذن.

التقدير الواقعي: وهو تقدير assessment الأداء بالاعتماد على مهام حقيقية ومستوحاة من الواقع.

سعة الطيف: وهو مصطلح أطلقه ل. ج. كرونباخ على مجموعة المحكات التي يمكن التنبؤ بها من خلال اختبارٍ ما، وكلما تعددت المحكات التي يمكن لاختبارٍ ما التنبؤ بها اتسع طيف هذا الاختبار. راجع مصطلح الأمانة.

العمر القاعدي: وهو أعلى مستوى عمري في اختبارات الذكاء (كالطبعات الأخيرة من اختبار ستانفورد - بينيه). يتجاوز المتحن وينجاح جميع الاختبارات الفرعية الأدنى منه أو المطابقة لمستواه.

النسبة القاعدية: وهي نسبة الأشخاص ضمن تعداد سكاني معين يمثلون خاصية، حالة أو سلوك معينين.

التحيز: وهو الاتجاه في العديد من المقاييس السياكومترية إلى التوصل لنتائج أعلى أو أدنى من النتائج الفعلية ؛ كذلك التحيز مع فئة من الناس على حساب فئة أخرى. راجع مصطلح العدالة.

السقف العمري: وهو أدنى مستوى عمري في اختبارات الذكاء (كالطبعات الأخيرة من اختبار ستانفورد - بينيه). يفشل المتحن في جميع اختباراته الفرعية.

التصنيف: وهو استخدام درجات الاختبار لوضع الأفراد ضمن مجموعات مختلفة.

تقنية كلوز: وهي إجراء اختباري يتم فيه مسح عدد من الكلمات من نص ما ثم يطلب من المتحن ملء فراغات النص. وتقييم قدرة المتحن على القراءة من خلال قدرته على فهم النص وملء الفراغات بالشكل المطلوب.

اختيار العينات: وهو إجراء لأخذ العينات يتم فيه تقسيم السكان إلى مجموعات أو أقسام. ويتناسب عدد الوحدات المختارة عشوائياً من كل مجموعة مع إجمالي عدد الوحدات في هذه المجموعة.

التدريب والإعداد: وهو إجراء تعليمي قصير الأمد يهدف إلى تحسين درجات المتحنين في الاختبارات التي سيخضعون لها في المستقبل. يتم تدريب المتحنين على

الاختبار مسبقاً من خلال التعرف على نمط الأسئلة والتدرب على استراتيجيات الاختبار.

المعامل ألفا: وهو معامل ثبات الإتساق الداخلي وهو مناسب للاستخدام في الاختبارات ذات الأسئلة التي تمنح درجة ١ أو ٠ والأسئلة ذات الدرجات المتعددة؛ كذلك معامل الإرتباط المتوقع لاختبار ما مع نظير له يحتوي عدداً مماثلاً من الفقرات أو الأسئلة.

معامل المطابقة Concordance: وهو معامل العلاقة بين الدرجات التي يمنحها مصحح ما لورقة اختبارية ما والدرجات التي يمنحها مصحح آخر لنفس الورقة.

معامل التكافؤ: وهو معامل للثبات (إرتباط) يمكن الحصول عليه عن طريق إجراء شكلين مختلفين من الاختبار ذاته للأفراد ذاتهم.راجع ثبات الصور المتكافئة.

معامل الإتساق الداخلي: وهو معامل للثبات مبني على تقديرات الإتساق الداخلي للاختبار؛ معاملات التجزئة النصفية، معاملات كودر - ريتشاردسون ومعامل ألفا.

معامل الإستقرار: وهو معامل للثبات يمكن الحصول عليه عن طريق إجراء الاختبار ذاته للأفراد ذاتهم في توقيتين مختلفين. راجع ثبات إعادة الاختبار.

معامل الإستقرار والتكافؤ: وهو معادل للثبات يمكن الحصول عليه عن طريق إجراء شكلين مختلفين من الاختبار ذاته للأفراد ذاتهم في توقيتين مختلفين.

الإدراك: وهو كل ما له علاقة بالفكر من ذاكرة وتفكير وحل المشكلات وغيرها.

التقدير الإدراكي: وهو مقياس للعمليات العقلية كالإدراك والذاكرة والتفكير والحكم والإستدلال.

النمط الإدراكي: وهو اتجاه أو نمط الإدراك والتذكر والتفكير الذي يحدده الفرد لنفسه بغرض فهم العالم من حوله والتأقلم معه (مثال: الاعتمادية والاستقلال، التعقل والاندفاعية، التحكم بالذات داخلياً أو تحكم العالم الخارجي بالفرد).

التشارك Communality: وهو نسبة التباين في متغير تم قياسه يمكن أن تعزى إلى التباين الذي يشترك بها هذا المتغيير مع غيره من المتغيرات.

الدرجات المركبة: وهي المجموع المباشر أو المثقّل لدرجات اختبارين أو أكثر أو قسمين من نفس الاختبار.

الاختبار التكيفي المحوسب: وهو إجراء اختباري عن طريق الحاسب تتكيف فيه الأسئلة والمهام اللاحقة في الاختبار تبعاً لإجابات المتحن على الأسئلة السابقة.

ثبات المطابقة: يقوم عدد من المصححين بإعطاء درجة عددية معينة لدرجة وجود سلوك أو خاصية ما عند عينة كبيرة من الناس، ثم يتم حساب معامل المطابقة (وهو مؤشر للتوافق بين آراء المصححين).

الصدق التلازمي: وهو درجة ارتباط الدرجات التي يحصل عليها مجموعة من الناس في اختبار سايكومتري معين مع الدرجات التي يمكن التوصل إليها عن طريق مقياس آخر للخاصية نفسها التي يحاول الاختبار قياسها.

مجال الثقة: وهو المجال الذي تقع ضمنه درجة الممتحن الحقيقية في اختبار ما أو في متغير المحك. ويتراوح مجال الثقة فيبين ٩٥٪ و٩٩٪. راجع الخطأ المعياري للتقدير والخطأ المعياري للقياس.

وزن الثقة: وهي طريقة موضوعية لوضع الدرجات يعتمد فيها الثقل العددي المعطى لإجابة صحيحة في اختبار ما على درجة ثقة الممتحن في صحة إجابته.

صدق البناء: وهي درجة ارتباط الدرجات التي يحصل عليها الفرد في اختبار سايكومتري مصمم لقياس خصائص معينة مع المقاييس التي تقيس السلوك في مواقف تحدد فيه الخاصية موضع القياس سلوك الفرد.

الفقرات الإنشائية: وهي أسئلة أو مسائل تتطلب إجابتها كتابة أو رسم بياني أو صورة أو غيرها مما يتطلب الإنشاء. ويعرف هذا النوع من الأسئلة أيضاً باسم "الأسئلة المفتوحة".

صدق المحتوى: وهو درجة تأكيد المختصين والخبراء على أن الاختبار أو المقياس السياكومتري يقيس في الواقع ما كان قد صمم لقياسه.

التصحيح من أثر التخمين: وهي معادلة تطبق على الدرجات الخام للاختبارات لتصحيح آثار الإجابات التي يخمنها الممتحن. ومن أكثر معادلات التصحيح هي تلك التي تطرح جزءاً من عدد الإجابات الصحيحة من عدد الإجابات الخاطئة.

الإرتباط: وهي درجة العلاقة بين متغيرين كالعلاقة بين الاختبار ومقياس المحك.

معامل الارتباط: وهو مؤشر عددي لدرجة العلاقة بين متغيرين ويتراوح عادة بين ١.٠٠- (علاقة سلبية كاملة) و ٠٠، (لا علاقة أبداً) إلى ١.٠٠+ (علاقة إيجابية كاملة). وهنالك نوعان من معامل الارتباط وهما المعامل المولد للعزوم ومعامل النقطة ثنائية التسلسل).

المحك: وهو مقياس أو متغير تُقاس أو تُقيم درجات أي اختبار سايكومتري بالإستناد إليه. ويُحدد صدق أي اختبار سايكومتري يُستخدم في إختيار أو تصنيف الأشخاص عن طريق قدرة هذا الاختبار على التنبؤ بمحك سلوكي معين مرتبط بالموقع الذي يتم إختيار الأشخاص له أو تصنيفهم ضمنه.

تشوه المحك: وهو تأثير أي عامل كان على المحك يؤدي إلى أن يصبح المحك غير صالح لقياس إنجازات الفرد. فعلى سبيل المثال تُستخدم نتائج اختبارات الاستعداد للتبؤ بدرجات التحصيل المدرسية ولكن استخدام المدرسين لنتائج اختبارات الاستعداد لتحديد درجات الطالب يجعل هذه الدرجات محكاً غير صالح لإثبات صدق اختبار الاستعداد وبالتالي فقد تم تشوه المحك.

الاختبار ذو المرجعية المحكية: وهو اختبار يُقيد محتواه بمواصفات خاصة ويخدم عدداً معيناً من الأغراض المحددة.وغالباً ما يكون الهدف من الاختبار هو تحديد درجة تحصيل المتحن بالمقارنة مع أهداف تعليمية معينة. راجع الاختبارات ذات المرجعية المعيارية.

الصدق المرتبط بالمحك: وهو درجة قدرة اختبار ما على قياس ما يريد هذا الاختبار أن يقيسه. ويمكن تحديد هذا الأمر بالنظر إلى الإرتباط بين درجات الاختبار ومقياس المحك لسلوك ما.

السلوك الحاسم Critical incident: وهو أي سلوك يتوقف عليه الأداء الفعال لفردٍ ما في وظيفة ما (مثال: "ينظف مكان العمل قبل المغادرة" و"يتعامل مع الزبائن بلباقة").

الصدق عبر المجموعات Cross validation: وهي طريقة لتحديد صدق اختبار ما مع مجموعة جديدة من الناس بعد أن تم إثبات صدق هذا الاختبار في التنبؤ بمحك

سلوكي ما عند جماعة بشرية أخرى. يتم هذا عن طريق إجراء هذا الاختبار للمجموعة الجديدة. وغالباً ما يتقلص معامل الصدق مع المجموعة الثانية من المتحنين إذ أن عامل الحظ يضخم معامل الصدق في المرة الأولى للاختبار.

القدرة المتبلورة (g): وهو مصطلح أطلقه ر. ب. كاتل على القدرات العقلية (القدرات المعرفية) المكتسبة من الخبرة والتعليم.

الاختبار العادل حضارياً: وهو اختبار يُفترض في مواده أن تكون متوفرة لجميع الشعوب والحضارات فلا يعاقب هذا الاختبار أية مجموعة حضارية بسبب قلة خبرتها في أمور معينة. وحتى الآن لم تكلل الجهود لتصميم اختبارات عادلة حضارياً بالنجاح.

الدرجات الفاصلة (الحد الأدنى للنجاح): يتم قبول جميع المتقدمين الذين حصلوا على درجات أكثر أو مساوية للدرجات الفاصلة ورفض المتقدمين الذين حصلوا على درجات دون هذا الحد . وتعتمد الدرجات الفاصلة للاختبار على صدق الاختبار، النسبة المحددة مسبقاً لعدد الناجحين بالنسبة لعدد المتقدمين، وغيرها من العوامل.

الدرجات المشتقة: وهو معدل للدرجات يتم التوصل إليه عن طريق إجراء عملية حسابية بسيطة على الدرجات الأولية للاختبار مثل ضرب معدل الدرجات الأولية في عدد ثابت أو/و إضافة هذا العدد الثابت للدرجات. راجع الدرجات المعيارية، درجات T، ودرجات z.

العمر التطوري (الارتقائي) (DA): وتمثله درجة التحصيل في جداول جيزل الإرتقائية.

معدل التطور (DQ) وهو مؤشر (DQ = 100 x DA/CA) يماثل بشكلٍ تقريبي العمر العقلي للرضع ويلخص سلوك الرضيع تبعاً لتقييم جداول جيزل.

معدل الذكاء الإنحرافي: وهو معدل لدرجات اختبار الذكاء يُستخلص عن طريق تحويل معدل الدرجات الخام لاختبار ذكاء إلى توزيع درجات وسطه الحسابي ١٠٠ وإنحرافه المعياري ثابت (فهو ١٦ لاختبار ستانفورد - بينيهو١٥ لاختبارات ويكسلر).

الاختبار التشخيصي: وهو اختبار تحصيلي مؤلف من عدد من المهارات التي تشكل مادة دراسية معينة يهدف إلى تحديد نقاط القوة والضعف لفرد ما في المجال

موضع الاختبار. ويتوفر هذا النوع من الاختبارات في مجالات القراءة والحساب والتهجئة.

عمل الفقرات التبايني (DIF): يحدث هذا عندما تتغير مؤشرات فقرات الاختبار عند تغير المجموعات موضع الاختبار (مثال:أن تجد إحدى المجموعات فقرات امتحانية ما أسهل من غيرها أو أكثر تحيزاً من غيرها).

صدق التمييز: وهي حالات وجود علاقات إرتباط ضعيفة بين أداة قياس سايكومترية ما ومقاييس أخرى للمفهوم العقلي.

المموه distracter: وهي اجابة خاطئة تماماً توضع ضمن إختيارات الإجابة للأسئلة المتعددة الإختيارات.

التقييم الديناميكي: وهو أسلوب في التقييم يعتمد على نموذج درّس- اختبر - درّس ويتم فيه اختبار الفرد (الاختبار الأولي) ثم تدريبه على مواد الاختبار ثم إعادة اختباره (الاختبار اللاحق). ويشكل اختلاف الأداء بين الاختبارين مقياساً لقابلية التعلم عند الفرد. راجع منطقة التطور الممكن.

مرض صعوبة الحساب (ديسكالكوليا): وهي عدم القدرة على إجراء العمليات الحسابية.

مرض صعوبة القراءة والكتابة (ديسليكسيا): وهو خلل في القدرة على القراءة يرتبط مع خلل القدرة على تفسير العلاقات المكانية أو الجمع بين المعلومات البصرية والسمعية.

المتخلفين عقلياً الذ ين يمكن تعليمهم (EMR): وهم الأطفال الذين تم تشخيص نسبة بسيطة من التخلف العقلي لديهم (وهم أولئك الذين تتراوح معدلات ذكائهم بين ٥١ و ٦٩). ويمكن لهؤلاء الأطفال أن يصلوا إلى مستوى تعليمي بين الصفين الثالث والسادس، كما يمكنهم أن يقوموا بعمليات حسابية بسيطة بالإضافة إلى القراءة والكتابة. راجع المتخلفين عقلياً الممكن تدريبهم.

الهدف التعليمي: وهو هدف أو برنامج تعليمي أو تأهيلي يحدد ما يجب على المتعلم أن يقدر على فعله مع نهاية الدرس أو الدورة التعليمية.

التصحيح عن طريق الاستبعاد: وهو إجراء لوضع الدرجات يقوم فيه المصحح بالإشارة إلى الإجابات الخاطئة لفقرة اختبارية ما بدلاً من تصحيح الإجابة الأفضل للفقرة فقط.

طريقة المؤشر (الخط) المئيني المعادل: وهي طريقة تقليدية لتحويل وحدات درجات إختبار ما إلى وحدات درجات لاختبار مكافئ. يتم تحويل درجات كلا الاختبارين إلى رتب مئينية وتوضع النتائج في جدول يتم التوصل إليه عن طريق معادلةالدرجات الواقعة على الخط المئيني p للاختبار الأول مع الدرجات الواقعة على الخط المئيني p للاختبار الثاني.

الصور الامتحانية المكافئة: راجع الشكل الامتحاني الموازي.

الاختبار المقالي: وهو نوع من الاختبارات يطلب فيه من الممتحنين الإجابة على عدد من الأسئلة بجواب واحد مطول. ويتم تقييم الإجابات من قبل المدرس أو المقيِّم بشكل ذاتي. راجع الاختبار الموضوعي.

جدول التوقع: وهو جدول يوضح تكرار أو نسبة الممتحنين الواقعين ضمن فئة معينة (حزمة درجات) من متغيِّر تنبؤي (اختبار) يتوقع أن يقعوا ضمن فئة معينة (فئات درجات) من متغيِّر المحك.

الصدق المظهري: وهي درجة كون مظهر أو محتوى الاختبار ومواده (أو فقراته أو ما شابه ذلك) تعطي الانطباع بأن الاختبار أو المقياس السايكومتري يقيس ما يجب أن يقيسه.

العامل factor: وهو بُعد أو صفة أوخاصية للشخصية يمكن الحصول عليها عن طريق تحليل عوامل مصفوفة الارتباط التي تم حسابها من درجات عدد كبير من الأفراد في اختبارات مختلفة أو فقرات من اختبارات مختلفة.

التحليل العاملي: وهو إجراء حسابي لتحليل مصفوفة الارتباط بين المقاييس لتحديد العوامل (أو المفاهيم) القادرة على تفسير الارتباط.

أحمال العامل Factors loading: هو، في التحليل العاملي، الارتباطات الناتجة (الأوزان) بين الاختبارات (أو غيرها من المتغيرات) والعوامل المستخلصة.

دوران العامل: Factor rotation وهو إجراء حسابي يتم إجراؤه على مصفوفة المعاملات بهدف تبسيطها لتفسير نتائجها. ويتم هذا عن طريق زيادة عدد تحميلات العاملين الأعلى والأدنى في المصفوفة. ويمكن أن يكون دوران العامل عامودياً بحيث تشكل المعاملات الناتجة زوايا قائمة مع بعضها، أو منحرفاً بحيث تشكل محاور العامل الناتج زوايا حادة أو منفرجة.

العدالة: وهي درجة عدم تحيز الدرجات النهائية في الاختبارات، أو كون هذه الدرجات تتنبأ بشكل متعادل بالأداء على المحك لمجموعات مختلفة من الأفراد.

الخطأ السلبي غير الحقيقي: وهو خطأ في الإختيار أو خطأ في القرار التشخيصي يتنبأ فيه الإجراء التقديري بنتيجة سيئة (مثال: تحصيل دون المستوى، أداء متدني، الاضطراب العقلي).

الخطأ الإيجابي غير الحقيقي: وهو خطأ في الإختيار أو خطأ في القرار التشخيصي يتنبأ فيه الإجراء التقديري بنتيجة جيدة (مثال: تحصيل أعلى من المستوى، أداء جيد، غياب الاضطراب العقلي).

الأمانة: وهو محدودية سعة طيف الاختبار أو غيره من أدوات القياس. ويتنبأ الاختبار ذو سعة الطيف المحدودة بشكل جيد بعدد محدد من المحكات. راجع سعة الطيف أو النطاق.

السؤال ذو البدائل المحدودة: وهو سؤال متعدد الاختيارات يتألف من جذر وعدة إجابات محتملة.

الاختبار ذو المستويات المرنة Fexilevel test: وهو اختبار يتألف من عدة أسئلة (فقرات) مرتبة من الأسهل إلى الأصعب ويبدأ الممتحن الاختبار من المنتصف وعندما يعطي الممتحن إجابة خاطئة عن سؤال ما يُطرح عليه سؤال أسهل، أما عندما يعطي إجابة صحيحة فيُطرح عليه سؤال أصعب.

القدرة الفطرية (g_f): وهي مصطلح أطلقه ر. ب. كاتل على القدرة العقلية الكامنة داخل الأفراد والتي تحددها الوراثة مثل قدرات حل المشاكل والاستجابات الإبداعية (غير التقليدية) للأمور.

التقدير التكويني: وهو طريقة لتقييم الأداء يتم عن طريقها إما تطوير أساليب التعليم والتدريس أو التعرف على نقاط الضعف والقوة لدى الطلاب بهدف تطوير مواد تعليمية تثري معرفة المتفوقين أو تحل المشاكل التعليمية للطلاب الضعيفين. راجع التقييم التراكمي.

قاعدة الأربعة أخماس: وهي قاعدة للاختيار تقضي بأن تدني نسبة معدل اختيار أي عرق أو جنس أو نوع في أي إجراء للاختيار دون نسبة الأربعة أخماس(٨٠٪) من العرق أو الجنس الذي له نسبة الاختيار الأعلى له آثار سلبية، وبالتالي فإجراء الاختيار هذا غير قانوني.

العامل g: وهو معامل عام للذكاء إفترضه سبيرمان لتفسير الترابط الشديد بين اختبارات الذكاء المختلفة.

معامل التعميم: وهو معامل عددي يعد مؤشراً على درجة القدرة على التعميم بالانتقال من العينة موضع الاختبار إلى السكان بشكل عام.

نظرية التعميم: وهي نظرية لدرجات الاختبارات وللتشكيل الإحصائي المرتبط بها والذي يعطي تصوراً لدرجات اختبار ما كعينة من ضمن مجموعة كبيرة من العينات. يستخدم تحليل إجراءات التباين، لتحديد القدرة على تعميم الدرجات كقيم عامة (كونية) لأداء المتحنين والفقرات الامتحانية والظروف الامتحانية.

"**القاعدة الذهبية**": وهي حل وسط توصلت إليه شركة تأمين القاعدة الذهبية وخدمة الاختبارات التعليمية (ETS) وافقت فيه ETS على إجراء اختبار لمندوبي التأمين مؤلف من فقرات اختبارية لا تحتوي على فروقات كبيرة بين العرق الأبيض والعرق الأسود.

معيار الصف Grade norm: وهو وسيط درجات الاختبار لعدة أطفال منتمين إلى سنة دراسية واحدة.

استراتيجية التلقين التدريجي Graduated Prompting Strategy: وهي استراتيجية للتقييم الديناميكي يقوم فيها مُجري الاختبار بتقديم عدد من التلميحات السلوكية بغرض تعليم المتحنين القواعد اللازمة لاجتياز الاختبار. وغالباً ما تكون التلميحات الموجودة في نص معد سابقاً والتي لا تعتمد على إجابات المتحنين عامة جداً أول الأمر ثم تصبح أكثر تحديداً.

الاختبار الجماعي: وهو أختبار يجريه فرد واحد لعدد من المتحنين في نفس الوقت.

التصحيح الشمولي: وهو إجراء لوضع الدرجات (وغالباً ما يستخدم في الاختبارات ذات الأسئلة المقالية) يتم فيه تقييم أداء المتحن بشكل عام ووضع درجات لإجاباته بناءً على ذلك دون أعطاء درجات مختلفة لعناصر الإجابة المختلفة. راجع وضع التصحيح التحليلي.

الصدق المتزايد: وهو زيادة في صدق الاختبار تنتج عن استخدام اختبار جديد في عملية الإنتقاء وتتجاوز هذه الزيادة صدق كل إجراءات الاختيار السابقة.

الاختبار الافرادي Individual test: وهو اختبار يُجرى لكل فرد على حدة.

الموافقة المطلعة: وهي الموافقة على اختبار نفسي أو المشاركة في دراسة نفسية بعد الإطلاع على جميع الحقائق أو المحاذير المتعلقة بالاختبار.

الذكاء: تتعدد تعريفات الذكاء حيث يرى أ. بينيه أنه القدرة على الحكم بشكل جيد والفهم بشكل جيد والمحاكمة بشكل جيد، ويعرّفه ل. م. تيرمان بالقدرة على التفكير المجرد، إلا أن اختبارات الذكاء بشكل عام تقيس قدرة الأفراد على النجاح في المهام المدرسية وشبيهاتها.

معدل الذكاء (IQ): وهو معدل درجات مشتق، استخدم أصلاً في وضع درجات على مقياس ستانفورد – بينيه للذكاء. يتم حساب معدل الذكاء النسبي بتقسيم عمر المتحن العقلي على عمره الفعلي ثم يُضرب الناتج ب ١٠٠، بينما يتم حساب معدل الذكاء الانحرافي بضرب الناتج z الموافق للدرجات الأولية لاختبار ذكاء ما بالعدد الذي يمثل الانحراف المعياري لمعدلات الذكاء الانحرافية ثم إضافة ١٠٠ للناتج.

اختبار الذكاء: وهو اختبار نفسي مصمم لقياس درجة استعداد الفرد للعمل المدرسي وغيره من النشاطات التي تدخل فيها القدرة اللغوية وقدرة حل المشكلات.

الفقرات الاختبارية المتداخلة: وهي فقرات اختبارية تعتمد أو تتأثر فيها الإجابة على أحد الأسئلة على إجابات الأسئلة الأخرى.

الاتساق الداخلي: وهو درجة كون جميع الفقرات الامتحانية قادرة على قياس المفهوم ذاته أو المتغير ذاته. وإن ثبات الاختبار الذي يقاس بواسطة معادلات سبيرمان - براون أو كودر - ريتشاردسون أو كرونباخ - ألفا هي مؤشر على الإتساق الداخلي للاختبار.

ثبات المصححين: وهي نوع من الثبات تحددها درجات للاختبار يضعها مصححان لاختبارات عينة من الأفراد.

المؤشر الفئوي: وهو مقياس يشير فيه تساوي الفروقات العددية إلى تساوي الإختلافات في الصفة أو الخاصية موضع القياس. وتعد مقاييس الحرارة (درجة مئوية وفهرنهايت) ومقاييس الدرجات القياسية مقاييس فئوية.

الفقرة الاختبارية: وهي إحدى الوحدات أو الأسئلة أو المهام التي تؤلف الاختبار السايكومتري.

تحليل الفقرات: وهو مصطلح عام يطلق على إجراء تقدير ثبات أو فائدة مجموعة ما من الفقرات الامتحانية.

التحيز في فقرات الاختبار Itembias: وهو درجة كون فقرة امتحانية ما تقيس مفاهيم مختلفة عند مجموعات عرقية أو حضارية أو إقليمية أو جنسية مختلفة.

منحنى خصائص الفقرة: وهو رسم بياني يستخدم في تحليل الفقرات الاختبارية يتم فيه رسم عدد المتحنين الذين اجتازوا فقرة ما بنجاح إلى جانب درجات الاختبار الكلية.

مؤشر صعوبة الفقرة: وهو مؤشر يدل على صعوبة أو سهولة فقرة اختبارية ما بالنسبة إلى عدد من المتحنين. ويتم قياس صعوبة فقرة ما عن طريق حساب النسبة المئوية (p) لعدد المتحنين الذين اختاروا الإجابة الصحيحة لهذه الفقرة.

مؤشر تمييز الفقرة: وهو مقياس لفعالية فقرة اختبارية ما في التمييز بين الأفراد الذين حصلوا على درجات عالية في الاختبار (أو في متغير آخر للمحك) وأولئك الذين حصلوا على درجات متدنية.

تقاطع الفقرات: وهو الدرجة التي تقيس فيها فقرتين مختلفين المتغير ذاته مما يؤدي إلى ارتباط درجات الفقرتين.

منحنى الإجابة على الفقرة: وهو رسم بياني يوضح عدد المتحنين الذين أجابوا على الفقرة بشكل صحيح يُرسم إلى جانب محك داخلي (مجموع درجات الاختبار) أو خارجي للأداء.

نظرية الإستجابة للفقرة IRT: وهي نظرية لفقرات الاختبار يتم فيها التعبير عن درجات الاختبارات على شكل درجات تقديرية ضمن مجال القدرة الكامنة.

الحصول على عينات الفقرات: وهو إجراء لإنتقاء فقرات لاختبارات فرعية من ضمن بنك للفقرات. وتقدم عينات مختلفة من الفقرات لمجموعات مختلفة من المتحنين.

تحليل الوظيفة: وهو مصطلح عام للإجراءات التي تحدد المهام المحددة التي تنطوي عليها وظيفة ما. ومن الضروري أن يتم هذا الإجراء قبل بناء اختبار للتنبؤ بالأداء الوظيفي للأفراد.

معادلات كودر - ريتشاردسون: وهي معادلات تستخدم في حساب ثبات الإتساق الداخلي لإجراء اختبار يُصحح على مبدأ ١-٠ (درجة واحدة للإجابة الصحيحة و٠ للإجابة الخاطئة) مرة واحدة.

اختبار اللغة: وهو اختبار مؤلف من فقرات لفظية وحسابية، أي فقرات تتطلب استخدام لغة ما (الأحرف أو الأرقام).

نظرية السمات الكامنة Latent trait theory: وهي واحدة من عدد من النظريات (كنظرية منحنى خصائص الفقرة ونموذج راش) مع الإجراءات الإحصائية المرافقة لها والتي تربط بين درجات الفقرات ودرجات الاختبارات من جهة وتقدير معين لموقع الفرد ضمن السمات أو المجال الإفتراضي للقدرة الكامنة من جهة أخرى؛ تُستخدم هذه النظرية في تحليل الفقرات وقياس الاختبارات.

الصعوبات التعليمية: وهي صعوبات يواجهها أفراد معدلات ذكائهم حول المعدل أو أعلى من المعدل في القراءة والحساب والكتابة والتهجئة وغيرها من المهارات الأكاديمية.

النعت الخالي من التجريح Least Stigmatizing label: وهي تصنيف أو وصف لحالة تم تشخيصها يعبر عن الحالة بشكل دقيق دون أن يجرح الأفراد المعنين أو أن يسبب لهم وصمة اجتماعية.

تحليل الإنحدار الخطي: وهو إجراء لتحديد المعادلة الجبرية للخط الأنسب للتنبؤ بقيمة متغير مستقل ما من متغير تابع أو أكثر.

الربط: وهو إجراء من نظرية IRT لمعادلة اختبارين عن طريق تحويل مؤشرات الفقرات لاختبار ما إلى مؤشرات للفقرات لصورة ثانية من نفس الاختبار الأمر الذي يضع المؤشرات المتوافقة للاختبارين على المقياس العددي ذاته.

المعايير المحلية: الرتب المئوية والدرجات القياسية وغيرها من المعايير الموافقة لدرجات اختبار أولية لمجموعة صغيرة من المتحنين المحليين.

اختبار البراعة: راجع الاختبار المبني على المحكات.

فقرات المطابقة: وهي نوع من فقرات الاختبار تطلب من المتحنين المطابقة بين جذور السؤال الموجودة في عمود أو قائمة ماوالإجابات الموجودة في قائمة (عمود) آخر.

القياس: وهو إجراء لحساب كمية أو مقدار مفهوم ما أو أي كيان آخر؛ كذلك تعيين أرقام للأحداث والأشياء.

الوساطة Mediation: وهي استراتيجية للتقدير الديناميكي يتفاعل فيها المتحن مع القائم على الاختبار بشكل متواصل وذلك بهدف زيادة احتمالية توصل المتحن إلى الحل المناسب لمسألة اختبارية ما.

العمر العقلي (MA): وهو معدل درجات مشتق لاختبار ذكاء كاختبار ستانفورد - بينيه. يتوافق العمر العقلي لمتحن ما مع العمر الفعلي (الزمني) لعينة من الأطفال تمثل المجتمع عمرها الفعلي واحد ومعدل درجاتها مساوٍ لدرجات هذا المتحن. راجع معدل الذكاء النسبي (IQ).

الوضع الدراسي المبني على العمر العقلي: وهو مؤشر للسنة الدراسية التي يعتبرفيها طالب ما فعال عقلياً.

الموهوبون عقلياً: أي شخص أداؤه الفكري أعلى بكثير من المعدل وغالباً ما يكون ذكاؤه بين ١٣٠ و١٤٠ أو أعلى.

المعاقون عقلياً: أي شخص أداؤه الفكري أدنى بكثير من المعدل وغالباً ما يكون ذكاؤه بين ٧٠ و٧٥ أو أدنى.

المتغير الوسيط: وهو متغير يؤثر على الإرتباط بين متغيرين آخرين.

الاختبار متعدد المستويات: وهو اختبار تم تصميمه ليناسب عدداً من المستويات والأعمار؛ كذلك اختبار منفصل يتم إعداده لكل مستوى على حدة.

بطارية الاستعدادات المتعددة: وهي بطارية اختبارات معايرة Conormed تم تصميمها لتقدير القدرات الفعلية.

فقرات الخيارات المتعددة: وهي فقرات اختبارية تتألف من جذر السؤال (تصريح أو سؤال أو عبارة أو غيرها) وعدد من الاختيارات أو البدائل واحد منها فقط صحيح.

معامل الإرتباط المتعدد (R): وهو مقياس لدرجة الإرتباط العام بين عدد من المتغيرات لها محك واحد للمتغير. يتراوح R بين + ١ و – ١ لعدة متغيرات واحد منها بعد متغيراً محكياً، ويتراوح الارتباط المتعدد العام لمجموعة من اختبارات الاستعداد مع معدلات الدرجات المدرسية بين ٦٠، و٧٠، وهي نسبة ارتباط معتدلة.

الدرجات الفاصلة المتعددة Multiple cutoft: طريقة لإنتقاء الموظفين أو الطلاب تعتمد على إحراز الممتحن لعدد من الحدود الدنيا للدرجات في عدد من محكات الإنتقاء ليتم قبوله في الوظيفة أو المؤسسة التعليمية.

تحليل الإنحدار المتعدد: وهو إجراء إحصائي لتحليل مساهمة متغيرين مستقلين أو أكثر في التنبؤ بقيمة متغير تابع.

المعايير الوطنية National norms: الرُتب المئينية والدرجات القياسية وغيرها من المعايير المبنية على عينة على مستوى الدولة. راجع: المعايير المحلية والمعايير.

التقدير العصبي – النفسي: وهو قياس الأداء الإدراكي والإدراكي – الحسي والحركي لتحديد موضع ودرجة وآثار الخلل الوظيفي العصبي والإصابات العصبية.

الإضطراب العصبي – النفسي: وهو اضطراب في الجهاز العصبي المركزي ترافقه أعراض نفسية.

المقياس الإسمي: وهو أدنى درجات القياس، وتستخدم فيه الأرقام كصفات أو أسماء للأشياء لا تعبر عن ترتيب أو كمية معينة.

الاختبار اللالفظي Nonverbal test: وهو اختبار لا يتطلب من الممتحن استخدام الكلمات المحكية أو المكتوبة بل استخدام الإشارة أو هز الرأس أو البناء أو التعامل مع الأشياء أو الإستجابة لأسئلة الاختبار بطريقة لا لفظية.

التوزيع الطبيعي: وهو توزيع سلس للدرجات يأخذ شكل الجرس. وهو متناظر حول الوسط الحسابي وتعبر عنه دالة رياضية دقيقة. وغالباً ما تتوزع درجات مجموعة كبيرة من المتحنين بشكل طبيعي.

الدرجات القياسية المطبّعة Normalized standard scores: وهي درجات قياسية (z) يتم الحصول عليها بتحويل درجات الاختبارات الأولية بطريقة تقود إلى توزيع هذه الدرجات توزيعاً طبيعياً.

مجموعة المعيار Norm group: عينة من الأفراد تم قياس الاختبار عليها.

الاختبار ذو المرجعية المعيارية: وهو اختبار يتم تفسير نتائجه بالإستناد إلى معايير تم الحصول عليها من عينة من الناس تمثل المجتمع. راجع الاختبار ذو المرجعية المحكية.

المعايير Norms: هي قائمة تضم معدلات درجات مع الرتب المئوية والدرجات القياسية وغيرها من الدرجات المشتقة لعينة قياس الاختبار الموافقة لهذه الدرجات.

الاختبار الموضوعي: وهو اختبار يتم تصحيحه (وضع درجاته) عن طريق مقارنة إجابات الممتحنين مع مفتاح للإجابات تم إعداده مسبقاً، على عكس الاختبارات الذاتية. ومن الأمثلة على الاختبارات الموضوعية اختبارات الخيارات المتعددة واختبارات الصواب والخطا.

الدوران المائل Oblique rotation: في التحليل العاملي، هو الدوران الذي تُشكل محاور العامل مع بعضها زوايا حادة أو منفرجة. ينتج عن ذلك الإرتباط بين العوامل.

ثبات الفقرات الفردية والزوجية: الإرتباط بين مجموع درجات الفقرات ذات الأرقام الزوجية ومجموع الدرجات الفقرات ذات الأرقام الفردية والذي يتم تصحيحه باستخدام معادلة سبيرمان – براون للثبات. راجع معادلة سبيرمان – براون.

كتيب الاختبار: وهو اختبار يتألف من عدد من الفقرات تقيس نواحٍ متعددة من النشاط العقلي. اختبار أوتيس – لينون للقدرة المدرسية واختبار هينمون – نلسون

للقدرة العقلية هي اختبارات تقدم على شكل كتيب اختباري. راجع كتيب الاختبار اللولبي.

الفقرات المفتوحة الإجابة: راجع الفقرات الإنشائية.

التعريف العملياتي: وهو تعريف يحدد ما يجب على فرد ما أن يعمله ليحس أو يجرب المفهوم المراد تعريفه.

العملياتية Operationalism: وهي عقيدة فلسفية تنطوي على التعبير عن مفهوم ما أو مصطلح علمي ما أو نظرية ما من خلال العمليات أو الوظائف التي تعرف أو تجسد هذا المفهوم أو المصطلح أو النظرية.

اختبارات المهارات اللغوية الشفهية (Oracy Test)(١): وهو اختبار لمهارات الإستماع والكلام.

الاختبار التحصيلي الشفهي: وهو اختبار يجيب فيه المتحنون شفهياً على أسئلة كتابية أو شفهية.

المقياس الرُتبي: وهو نوع من أنواع القياس تعبر فيه أرقام الأشياء عن رتب أو مراتب الأحداث والأشياء (كترتيب الإنتهاء في مسابقة جري أو غيرها).

الدوران العمودي Orthogonal rotation: في التحليل العاملي، وهو تدوير يحافظ على استقلالية العامل أي يُبقي على الزوايا القائمة بين العوامل مما يعني عدم وجود أي ارتباط بين هذه العوامل.

الاختبار الخارج عن المستوى: وهو اختبار يناسب عمر أو سنة دراسية أدنى أو أعلى من عمر المتحن أو سنته الدراسية.

الصور المتكافئة: وهما اختباران تتشابه فقراتهما من حيث درجة الصعوبة وترتفع نسبة الإرتباط بينهما. كما تتقارب الدرجات التي يحصل عليها الأفراد في الاختبارين بشكل كبير.

---

(١) كلمة oracy هي كلمة جديدة في اللغة الإنجليزية لم تتم إضافتها إلى معظم المعاجم حتى الآن وهي كلمة مركبة على نمط كلمة literacy التي تشير إلى مهارات الكتابة والقراءة. ومن الملاحظ أن هذه الكلمة تجمع بين كلمتي oral ولاحقة cy - (المترجم).

ثبات الصور المكافئة: وهو مؤشر لثبات الاختبار يُحدد عن طريق حساب الإرتباط بين درجات صورتين متشابهتين من الاختبار ذاته.

تحليل المسار: وهو إجراء إحصائي يتم فيه تحليل العلاقات بين المتغيرات لتحديد تسلسلها أو تتابعها السببي.

الخط المئيني: وهو معدل لدرجات اختبار ما تقع النسبة المئوية لمعدلات درجات معينة تحته أو عليه.

الحزمة المئينية: وهي عدد من الرتب المئينية تشير إحتماليات محددة إلى وقوع الدرجات الفعلية لممتحن ما ضمنها.

المعايير المئينية: وهي قائمة بالدرجات الخام والنسب المئوية الموافقة لها الناتجة عن عينة القياس والتي تقع تحت الخط المئيني المعطى.

الرتب المئينية: وهي النسب المئوية للدرجات التي تقع تحت معدل درجات معين في التوزيع التكراري للدرجات؛ كذلك النسبة المئوية الموافقة لمعدل درجات معين.

التقدير الأدائي: وهو نوع من التقييم يعتمد على البناء أو الإبداع أو العرض والتوضيح. وتتوفر في معظم الحالات عدة طرق للتقدير وأكثر من طريقة واحدة للإجابة بشكل صحيح.

اختبار الأداء: وهو اختبار يطلب من الممتحنين التعامل مع عدد من الأشياء الملموسة؛ تعتبر هذه الاختبارات نقيض اختبارات الورقة والقلم (أو الاختبارات الموضوعية).

معامل نقطة الارتباط ثنائي التسلسل Point-biserial coefficient: وهو معامل للإرتباط بين متغير يحتمل احتمالين أو قيمتين فقط (dichotomous) ومتغير مستمر يحتمل عدد من الاحتمالات (continuous)[٢].

مقياس النقاط Point scale: وهو اختبار يتم فيه إعطاء درجات لإجابات الطلاب (٠ أو ١ أو ٢... إلخ) بناءً على صحة الإجابة وشموليتها.

---

(٢) بعض المتغيرات كالجنس لا يحتمل إلا متغيرين (ذكر أو أنثى) اما العمر أو معدل الذكاء فهي أمثلة على المتغيرات المتعددة الإحتمالات، (المترجم).

ملف إنجاز الطالب (البورتفوليو): وهو مجموعة من أعمال الطالب الأدائية أو الإنتاجية يتم تجميعها خلال فترة زمنية معينة بغرض تقييمها.

اختبار القوة: وهو اختبار يُعطى فيه المتحنون وقتاً طويلاً للإجابة على أسئلة الاختبار التي قد تكون على درجة عالية من الصعوبة.

الصدق التنبؤي: وهي درجة قدرة درجات اختبار ما على التنبؤ بالأداء في محكٍ ما يتم قياسه لاحقاً؛ يتم التعبير عن الصدق التنبؤي على شكل إرتباط بين الاختبار (متغير التنبؤ) ومتغير المحك.

خصوصية التواصل Privileged communication: وهي أي تبادل للمعلومات بين الفرد ومحاميه أو طبيبه أو القسيس أو زوجه/زوجته يُمنح فيه مُعطي المعلومات الحق في منع الطرف الثاني من الإدلاء بشهادة رسمية تكشف أي من هذه المعلومات.

الاختبار التشخيصي: وهو اختبار يستخدم للتنبؤ بقدرة فرد ما في مادة معينة (كاختبار الجاهزية للقراءة).

المهارات الحركية - النفسية: وهي المهارات التي تتطلب حركات معينة كرمي الكرة والإمساك بها وإدخال شيء في شيء آخر وغيرها من الحركات التي تسمح للفرد بالتعامل مع الأشياء (المهارات الرياضية).

المعايرة العرقية Race norming: وضع درجات معيارية للاختبارات تستند إلى عينة من عرق أو فئة إثنية معينة، وتقييم المتقدمين بناءً على العرق الذي ينتمون إليه.

العينة العشوائية: وهي عينة من الملاحظات (كدرجات الاختبارات) يتم الحصول عليها من عينة من المجتمع تم اختيارها بطريقة تضمن لكل فرد من العينة الفرصة نفسها في الاختيار.

المدى (المجال): وهو مقياس خام لحدود أو تنوع مجموعة من درجات الاختبار يتم حسابه عن طريق طرح معدل الدرجات الأدنى من معدل الدرجات الأعلى.

التواصل: وهو علاقة ودودة وحميمة تنشأ بين الممتَحن والقائم على الاختبار.

نموذج راش: وهو نموذج أحادي المؤشر لحساب صعوبة الفقرات يستخدم في تحليل فقرات الاختبار وقياس الاختبارات. يستند هذا النموذج إلى فرضية أن مؤشرات

التخمين وتمييز الفقرة هي مؤشرات لا تؤخذ بعين الاعتبار نظراً لصغر قيمها. ويعتمد نموذج راش، شأنه كشأن باقي نماذج السمات الكامنة، على تحديد موقع الطالب في المجال الإفتراضي للقدرة الكامنة بناءً على أدائه في فقرات الاختبار المختلفة.

معدل الذكاء النِسبي Ratio IQ: وهو معدل للذكاء يتم الحصول عليه بتقسيم درجات العمر العقلي لفرد ما في أحد اختبارات الذكاء (كالنسخ الأقدم من اختبار ستانفورد - بينيه) على عمره الفعلي ثم ضرب الناتج بالعدد ١٠٠، راجع معدل الذكاء الإنحرافي.

المقياس النِسبي: وهو مؤشر للقياس قيمته الدنيا هي الصفر وتدل النسب العددية المتساوية بين فواصله على نسب متساوية للصفة أو الخاصية موضع القياس. وعلى حين لا يتم عادةً قياس المتغيرات النفسية باستخدام المؤشر النسبي، يتم استخدامه لقياس متغيرات أخرى كالطول والعرض والوزن وغيرها من المتغيرات الفيزيائية.

اختبار الجاهزية: وهو اختبار يقيس مهارات ومعرفة فرد ما لتحديد جاهزيته لتعلم المواد الدراسية المعقدة كالقراءة والرياضيات والتهجئة.

فقرات إعادة الترتيب: وهي نوع من الفقرات الاختبارية التي تتطلب من الممتحن إعادة ترتيب أجزاء السؤال وفقاً لتسلسلها الصحيح.

معادلة الإنحدار: وهي معادلة خطية للتنبؤ بدرجات المحك من درجات متغير تنبؤي واحد أو أكثر؛ كذلك إجراء يستخدم في عمليات الإنتقاء (للطلاب والموظفين) أو في حساب المخاطر أو في عمليات التشخيص.

الإنحدار نحو الوسط الحسابي: وهي نزعة لدى درجات الاختبارات وغيرها من المقاييس النفسية في الاقتراب من الوسط الحسابي عند إعادة الاختبار؛ ويزداد اقتراب معدل درجات ما من الوسط الحسابي عند إعادة الاختبار عندما يكون معدل الدرجات الأصلي شديد التطرف.

الثبات: الحد الذي بالنسبة له تقيس أداة تقدير نفسي أي شيء باتساق. وغالباً ما تتميز الدرجات الناتجة عن المقاييس الثابتة بأنها قريبة من حيث القيمة العددية من الدرجات الحقيقية للممتحنين وخالية نسبياً من الأخطاء.

مـعـامل الثـبـات: وهو مـؤشـر عـددي بين ٠,٠٠ و ١,٠٠ يعـبـر عن ثبـات أدوات القياس. يتم حساب ثبات الاختبارات باستخدام عدة طرق كطريقة إعادة الاختبار والصور المتكافئة والإتساق الداخلي.

قواعد التقييم (rubrics): وهو إجراء أو مخطط يتألف من عدد من المعايير يتم تقييم أداء المتحنين على اساسها.

اختبار الاستعداد المدرسي: وهو اختبار يتنبأ بقدرات الأفراد على تلقي المعلومات التي تُدرّس على مستوى المدرسة، ويشابه هذا النوع من الاختبارات اختبارات الذكاء التقليدية.

المسح: هو مصطلح عام يستعمل لأي عملية انتقاء ويكون عادة غير دقيق يتم فيه الاحتفاظ ببعض المتقدمين والإستغناء عن الآخرين.

الاختبار الآمن: وهو اختبار يتم ضمن إجراءات أمنية مشددة يضمن عدم اختبار أي فرد غير الأفراد المعنيين وعدم نقل أوراق الإجابة خارج قاعة الإمتحان حتى من قبل القائمين على الاختبار.

الإنتقاء: وهو عملية استخدام عدد من الاختبارات لاختيار المتقدمين الذين ترتفع احتمالية نجاحهم في المقرر أو التخصص أو المنصب الذي تقدموا له. ويتم قبول الأفراد الذين تقع معدلات درجاتهم فوق أو على الحد الفاصل للدرجات، بينما يتم رفض كل من تقع درجاته دون هذا الحد.

نسبة (معدل) الانتقاء: وهو نسبة المتقدمين الذين تم انتقائهم للإلتحاق بمقرر أو برنامج تدريبي أو عمل ما.

المعالجة التلاحقية (التتابعية) للمعلومات: عملية عقلية تتم فيها معالجة المعلومات بشكل متتابع أو متلاحق. ومن اللأمثلة على المعالجة التلاحقية محاولة تذكر سلسلة من الأرقام. راجع المعالجة المتزامنة.

الاختبار التتابعي: وهو نوع من الاختبارات تعتمد فيه الأسئلة التي تعطى للممتحن على إجاباته السابقة. يطلق على هذا النوع من الاختبارات كذلك اسم الاختبارات التكيفية أو الاختبارات المكيّفة.

الفقرات ذات الإجابات القصيرة: وهي فقرات لا تتطلب إلا إجابة قصيرة لملء الفراغ أو الإجابة على السؤال.

المعالجة المتزامنة للمعلومات: عملية عقلية يتم فيها التأليف بين أجزاء متفرقة من المعلومات في آنٍ واحد. راجع المعالجة التلاحقية.

مخطط سنيلين: وهو لوحة مؤلفة من أحرف صغيرة وكبيرة تقيس دقة وحدة البصر عند الأفراد.

معادلة سبيرمان - براون: وهي معادلة عامة لتقدير ثبات اختبار ما (r_{11}) تم زيادة فقراته بمقدار m. ففي المعادلة ، $r_{mm} = mr_{11}/[1+r_{11}(m-1)]$، حيث r_{11}، هي ثبات الاختبار الأصلي (قبل الزيادة) و m هي العدد المضاف.

الأطفال ذوي الخصوصية (الإحتياجات الخاصة): وهم الأطفال الذين يعانون من مشاكل جسدية أو نفسية أو إدراكية تجعل عملية وصولهم إلى أهدافهم وتلبية حاجاتهم الخاصة أصعب مما هي عليه مع غيرهم من الأطفال.

أدوات التحديد: وهي ألفاظ مثل: كافة ودائماً وإطلاقاً وفقط تشير إلى أن سؤال صواب أو خطأ ما هو خطأ على الأرجح، أو ألفاظ مثل: غالباً وأحياناً وعادةً تشير إلى أن سؤال صواب أو خطأ ما هو صحيح على الأرجح.

الإعاقات التعليمية الخاصة: راجع الإعاقات التعليمية.

الخصوصية: وهو الجزء من تباين الاختبار الذي يُعزى إلى عوامل خاصة بهذا الاختبار.

اختبار السرعة: وهو اختبار مؤلف من عدد كبير جداً من الأسئلة السهلة نسبياً ووقته محدود جداً مما يؤدي إلى فشل معظم المتحنين في إتمام الإجابة على كافة الأسئلة ضمن المدة الزمنية المتاحة. وغالباً ما تكون اختبارات المهارات الميكانيكية والمكتبية والحركية - النفسية اختبارات سرعة.

اختبار الكتيب اللولبي: وهو نوع من الاختبارات يتم فيه ترتيب الفقرات من الأسهل إلى الأصعب على شكل لولبي مما يعني أن الفقرات المتشابهة تتوزع على مراحل الاختبار المختلفة وتأتي مع الفقرات المماثلة في الصعوبة وليس في المحتوى.

معامل التجزئة النصفية: وهو طريقة لحساب ثبات الاختبار تعتمد على تطبيق معادلة سبيرمان - براون (m = 2) على الإرتباط بين نصفي الاختبار ذاته (كالفقرات ذات الأرقام الزوجية والفقرات ذات الأرقام الفردية).

الإنحراف المعياري: وهو الجذر التربيعي للتباين؛ يستخدم لقياس إنتشار مجموعة من معدلات الدرجات.

الخطأ المعياري للتقدير: وهو مقدار الخطأ في تقدير درجات فرد ما في متغير المحك بالإستناد إلى درجاته في متغير التنبؤ.

الخطأ المعياري للقياس: وهو تقدير للإنحراف المعياري للتوزيع الطبيعي للدرجات التي يُفترض أن يحصل عليها فرد ما إذا ما تم اختباره باستخدام ذات الاختبار عدداً لا نهائياً من المرات. إذا كانت الدرجات التحصيلية لفرد ما هي X فإن احتمال وقوع درجات هذا الفرد ضمن مجموعة الأفراد الذين تقع درجاتهم الحقيقية على بعد خطأ معياري واحد من Xهو واحد إلى ثلاثة.

الدرجات القياسية: وهي مجموعة من الدرجات كدرجات z ودرجات T أو درجات ودرجات الرتب من تسعة (STANINE) تمتلك الوسط الحسابي والإنحراف المعياري المطلوب. يتم حساب الدرجات القياسية عن طريق تحويل الدرجات الأولية إلى درجات z، ثم ضرب درجات z بالإنحراف المعياري المطلوب، ثم إضافة الوسط الحسابي المطلوب للدرجات المحوّلة إلى الناتج.

القياس: وهو إجراء اختبار تم بناؤه بعناية فائقة لمجموعة كبيرة من الأفراد تمثل المجتمع تحت ظروف ثابتة وذلك بغرض تحديد معايير للاختبار.

عينة القياس: وهي مجموعة فرعية من المجتمع الذي يستهدفه الاختبار يتم قياس الاختبار على أساسها.

الاختبارات القياسية: وهي اختبارات تم إعدادها بعناية ودقة من قبل محترفين، ثم تم إجراؤها تحت ظروف مثالية وتبعاً لتعليمات ثابتة لمجموعة من الأفراد تمثل المجتمع بغرض التوصل إلى معايير هذه الاختبارات.

الرتب من تسعة (STANINE): مقياس للدرجات المعيارية يتألف من الأعداد ١ إلى ٩ وسطه الحسابي هو ٥ وإنحرافه المعياري هو ٢ تقريباً.

اختيار العينات الطبقي العشوائي: وهو إجراء لأخذ عينات يعتمد على تقسيم المجتمع إلى طبقات (كالنساء والرجال، أو البيض والسود، أو الطبقة الفقيرة والطبقة المتوسطة والطبقة الغنية) ثم إنتقاء عينات عشوائية من كل طبقة مع مراعاة تناسب عدد الأفراد الذين تم اختيارهم من كل طبقة مع حجم الطبقة التي ينتمون إليها.

التقييم التراكمي Summative evaluation: وهو نوع من التقييم يتم في نهاية المقرر أو الوحدة الدرسية ويهدف إلى إعطاء المجموع العام أو القياس النهائي لتحصيل الأفراد.

الاختبار المسّاح: وهو اختبار تحصيلي يخالف الاختبار التشخيصي في خصائصه ويهدف إلى تقييم الأداء العام للممتحنين.

درجات T: وهي درجات مشتقة (محوّلة) قياسية مطبّعة وسطها الحسابي ٥٠ وإنحرافها المعياري ١٠. وعلى حين تشابه درجات z درجات T من حيث كون وسطها الحسابي ٥٠ وإنحرافها المعياري ١٠، تبقى درجات z درجات غير مطبّعة.

جدول المواصفات: وهو جدول ذو اتجاهين يتم إعداده كمخطط أو كإطار هيكلي للاختبارات التحصيلية، ويتم فيه تحديد الأهداف السلوكية في الحقول الأفقية وأهداف المحتوى (الموضوع) في الحقول العمودية.

الاختبار المكيّف: راجع الاختبار التتابعي.

جداول تايلر - راسل: وهي جداول تقييم لصدق الاختبارات على أساس دلالة المعلومات التي يتم الحصول عليها من الاختبار والتي لا يلعب فيها الحظ أي دور.

تصنيف الأهداف التعليمية: وهي مجموعة من الأهداف المرتبطة بدرس أو اختبار تحصيلي ما مرتبة طبقياً تشمل الطبقات الأعلى فيها كل ما ينتمي إلى الطبقات الأدنى.

تصنيفات السلوك الحركي - النفسي: وهي تصنيفات لأهداف درس أو اختبار ما ضمن المجال الحركي - النفسي.

الاختبار: وهو أية أداة تقيم السلوك أو الأداء لدى الأفراد. وتضم الاختبارات النفسية الاختبارات الإدراكية والعاطفية والحركية - النفسية وغيرها.

القلق المرتبط بالاختبارات Test anxiety: وهو الشعور بالقلق الذي ينتاب الأفراد وقت الاختبار.

التحيز في الاختبارات: يوجد هذا في الاختبارات التي تميز بشكل غير عادل بين مجموعتين أو أكثر من الناس.

ثبات إعادة الاختبار: طريقة لحساب ثبات الاختبار تعتمد على إجراء الاختبار ذاته لنفس المجموعة من الأفراد في وقتين مختلفين ثم حساب الترابط بين درجات الأفراد في المرتين.

البراعة في الاختبارات: وهي عدد من الطرق والإستراتيجيات التي يعتمدها بعض الممتحنين لتحسين درجاتهم في الاختبارات والتي لا ترتبط بمعرفتهم بمادة الاختبار.

الاختبار التجاري: وهو اختبار يقيس معرفة ومهارات الممتحن في مجال مرتبط بالوظائف والأعمال المختلفة؛ يُستخدم هذا النوع من الاختبارات في عمليات إنتقاء وتنسيب الموظفين ومنح التراخيص المهنية.

المتخلفين عقلياً الممكن تدريبهم (TMR): وهم الأطفال الذين تعتبر درجة تخلفهم العقلي (تتراوح معدلات ذكاء هؤلاء الأطفال بين ٣٦ و٥٠) درجة لا يمكن معها تعلم القراءة والكتابة إلا أنها تسمح بأداء المهام التي لا تحتاج إلى مهارة تحت المراقبة.

فقرات الصواب والخطأ: وهي نوع من الفقرات الاختبارية تحتوي على بيانات قد تكون صحيحة أو خاطئة.

الدرجات الفعلية (الحقيقية): معدل افتراضي لدرجات فرد ما يعبر عن معرفة الفرد الفعلية لمادة الاختبار. تنص نظرية الاختبارات على أن حساب الدرجات الفعلية للفرد يتم عن طريق حساب الوسط الحسابي لتوزيع درجات هذا الفرد إذا ما تم اختباره باستخدام الاختبار ذاته عدداً لا نهائياً من المرات.

الصدق: وهو درجة قياس أداة قياس ما للمتغير الذي تم تصميمها لقياسه. ومن أنواع صدق الاختبارات: صدق المحتوى (ويتم حسابه عن طريق تحليل محتوى الاختبار)، والصدق التنبؤي والتلازمي (ويتم حسابه عن طريق الربط بين درجات

الاختبار ودرجات مقياس المحك)، وصدق المفهوم (ويتم حسابه عن طريق دراسة مستفيضة لمدى قياس اختبار ما لمفهوم نفسي ما).

معامل الصدق: وهو الارتباط بين درجات المتغير التنبؤي ودرجات متغير المحك.

تعميم الصدق: وهو تطبيق أدلة الصدق على مواقف غير تلك التي تم منها الحصول على هذه الأدلة.

التغاير (التنوع): وهو توزع أو انحراف عدد من الدرجات حول قيمها الوسطية.

التباين: وهو مقياس لتغاير (تنوع) درجات اختبار ما يتم حسابه كمجموع مربعات إنحرافات الدرجات الخام من الوسط الحسابي مقسماً على عدد ينقص عن مجموع عدد معدلات الدرجات بمقدار ١؛ كذلك مربع الإنحراف المعياري.

القدرة اللغوية: القدرة على فهم الكلمات والعلاقات بينها في لغة ما؛ كذلك القدرة على حل المسائل المرتبطة بالكلمات.

الاختبار اللغوي (اللفظي): وهو اختبار تعليماته لغوية وتتطلب الإجابة الشفهية أو الكتابية على أسئلته كلمات أو أعداد.

قائمة الرصد للجماعات: وهي قائمة رصد يستخدمها المدرس لتقييم أداء مجموعة من الطلاب في آنٍ واحد.

اختبار عينة العمل: وهو اختبار تشابه فقراته المهام التي سيقوم بها المتحن في حال نجاحه في الحصول على الوظيفة التي تقدم إليها بطلب توظيف.

درجات z: أي من مجموعة من الدرجات المشتقة تتراوح بين ∞ – و ∞ + ويتم حسابها من المعادلة: z = (الدرجات الخام – الوسط الحسابي)/ الانحراف المعياري، وذلك لكل معدلات الدرجات الخام. وفي التوزيع الطبيعي للدرجات، تقع ٩٩٪ من الحالات بين z = –٣ و z = +٣.

منطقة التطور الممكن: وهي المسافة أو الفرق بين المستوى التطوري الفعلي للطفل (وهي مستوى التطور الذي حققه الطفل والذي تم تقديره عن طريق اختبار قياسي) وإمكانية تطوره (وهي درجة الكفاءة التي يمكن أن يصل إليها الطفل مع المساعدة).

المراجع

Aiken, L. R. (1970). Scoring for partial knowledge on the generalized rearrangement test item. *Educational and Psychological Measurement, 30*, 87–94.

Aiken, L. R. (1983). Determining grade boundaries on classroom tests. *Educational and Psychological Measurement, 43*, 759–762.

Aiken, L. R. (1987). Testing with multiple-choice items. *Journal for Research and Development in Education, 20*(4), 44–58.

Aiken, L. R. (1988). A program for computing the reliability and maximum reliability of a weighted composite. *Educational and Psychological Measurement, 48*, 703–706.

Aiken, L. R. (1989a). Some simple item analysis statistics and significance tests. *Journal for Research and Development in Education, 22*(2), 42–48.

Aiken, L. R. (1989b). Analyzing optional test items. *Educational and Psychological Measurement, 49*, 783–787.

Aiken, L. R. (1991). Detecting, understanding, and controlling for cheating on tests. *Research in Higher Education, 32*, 725–736.

Aiken, L. R. (1996). *Rating scales and checklists: Evaluating behavior, personality and attitudes.* New York: Wiley.

Aiken, L. R. (1997a). *Psychological testing and assessment* (9th ed.). Needham Heights, MA: Allyn & Bacon.

Aiken, L. R. (1997b). *Questionnaires and inventories: Surveying opinions and assessing personality.* New York: Wiley.

Airasian, P. W., & Terrasi, S. (1994). Test administration. In T. Husén & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (2nd ed.) (Vol. 11, pp. 6311–6315). Tarrytown, NY: Elsevier.

Allison, D. E. (1984). The effect of item-difficulty sequence, intelligence, and sex on test performance, reliability, and item difficulty and discrimination. *Measurement and Evaluation in Guidance, 16*, 211–217.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1985). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47*, 1597–1611.

American Psychological Association, Committee on Professional Standards, & Committee on Psychological Tests and Assessment. (1986). *Guidelines for computer-based tests and interpretations.* Washington, DC: American Psychological Association.

Anastasi, A. (1982). *Psychological testing* (5th ed.). New York: Macmillan.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). New York: Macmillan.

Angoff, W. H., & Ford, S. F. (1973). Item-race interaction on a test of scholastic aptitude. *Journal of Educational Measurement, 10,* 95–106.

Bender, L. (1938). *A visual motor gestalt test and its clinical use* (Research Monograph No. 3). New York: American Orthopsychiatric Association.

Binet, A., & Simon, T. (1905). Methodes novelles pour le diagnostic due niveau intellectuel des anormaux. *L'Année Psychologique, 11,* 191–244.

Black, H. (1962). *They shall not pass.* New York: Morrow.

Blakley, B. R., Quinones, M. A., Crawford, M. S., & Jago, I. A. (1994). The validity of isometric strength tests. *Personnel Psychology, 47,* 247–274.

Bloom, B. S., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: Handbook I. The cognitive domain.* New York: McKay.

Bloom, B. S., & Krathwohl, D. R. (1984). *Taxonomy of educational objectives: Handbook I. Cognitive domain.* White Plains, NY: Longman.

Bredemeier, M. (1991, November/December). IQ test ban for blacks called unconstitutional. *California Association of School Psychologists Today,* 22–23.

Bridgman, P. W. (1927). *The logic of modern physics.* New York: Macmillan.

Brigham, C. C. (1923). *A study of American intelligence.* Princeton, NJ: Princeton University Press.

Brigham, C. C. (1930). Intelligence tests of immigrant groups. *Psychological Review, 37,* 158–165.

Budoff, M. (1987). The validity of learning potential assessment. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 52–81). New York: Guilford Press.

Camilli, G. (1979). *A critique of the chi square method for assessing item bias.* Unpublished paper, Laboratory of Educational Research, University of Colorado, Boulder.

Campione, J. C., & Brown, A. L. (1987). Learning ability and transfer propensity as sources of individual differences in intelligence. In P. H. Brooks, R. D. Sperber, & C. McCauley (Eds.), *Learning and cognition in the mentally retarded* (pp. 265–294). Baltimore: University Park Press.

Carey, L. M. (1994). *Measuring and evaluating school learning* (2nd ed.). Needham Heights, MA: Allyn & Bacon.

Carroll, J. B. (1973). The aptitude-achievement distinction: The case of foreign language aptitude and proficiency. In D. R. Green (Ed.), *The aptitude-achievement distinction.* Monterey, CA: CTB/McGraw-Hill.

Carroll, J. B. (1978). On the theory-practice interface in the measurement of intellectual abilities. In P. Suppes (Ed.), *Impact of research on education* (pp. 1–105). Washington, DC: National Academy of Education.

Cattell, J. M. (1890). Mental tests and measurements. *Mind, 15,* 373–381.

Cohen, N. J. (1984). Preserved learning capacity in amnesia: Evidence for multiple sensory systems. In L. R. Squire & N. Butters (Eds.), *Neuropsychology of memory.* New York: Guilford Press.

Collet, L. S. (1971). Elimination scoring: An empirical evaluation. *Journal of Educational Measurement, 8,* 209–214.

Conoley, J. C., & Impara, J. C. (Eds.). (1995). *The twelfth mental measurements yearbook.* Lincoln: Buros Institute of Mental Measurements of the University of Nebraska-Lincoln.

Conoley, J. C., & Kramer, J. J. (Eds.). (1989). *The tenth mental measurements yearbook.* Lincoln: Buros Institute of Mental Measurements of the University of Nebraska-Lincoln.

Cooley, W. W., & Lohnes, P. R. (1968). *Predicting development of young adults* (Project TALENT Five-year Follow-up Studies, Interim Report 5, American Institutes for Research). Washington, DC: U.S. Office of Education.

Crowl, T. K., & McGinitie, W. H. (1974). The influence of students' speech characteristics on teachers' evaluations of oral answers. *Journal of Educational Psychology, 66,* 304–308.

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence.* Needham Heights, MA: Allyn & Bacon.

David, P. (1981, November 27). Multiple-choice tests under fire: United States. *Times Education Supplement, 3413*(11).

Davis, C. J. (1980). *Perkins-Binet tests of intelligence for the blind.* Watertown, MA: Perkins School for the Blind.

Delahunty, R. J. (1988). Perspectives on within-group scoring. *Journal of Vocational Behavior, 33,* 463–477.

Ebel, R. L. (1979). *Essentials of educational measurement* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Edouard, L. J., & Harris, F. T. C. (1983). Multiple-choice questions: Fact, fallacy and fancy. *International Journal of Mathematical Education in Science and Technology, 14,* 685–691.

Egeland, B. (1985). Review of Wisconsin card sorting test. In J. V. Mitchell (Ed.), *The ninth mental measurements yearbook* (pp. 1746–1747). Lincoln: University of Nebraska Press.

Embretson, S. E. (1991). A multidimensional latent trait model for measuring learning and change. *Psychometrika, 49,* 175–186.

Evans, W. (1984). Test wiseness: An examination of cue-using strategies. *Journal of Experimental Education, 52,* 141–144.

Eyde, L., Robertson, G. J., Krug, S., & Others. (1993). *Responsible test use: Case studies for assessing human behavior.* Washington, DC: American Psychological Association.

Fabiano, E. (1989). *Index to tests used in educational dissertations.* Phoenix, AZ: Oryx Press.

Feuerstein, R., Rand, Y., Jensen, M. R., Kaniel, S., & Tzuriel, D. (1987). Prerequisites for assessment of learning potential: The LPAD model. In C. S. Lidz (Ed.), *Dynamic assessment: An interactive approach to evaluating learning potential* (pp. 35–51). New York: Guilford Press.

Fremer, J., Diamond, E. E., & Camara, W. J. (1989). Developing a code of fair testing practices in education. *American Psychologists, 44,* 1062–1067.

Galton, F. (1883). *Inquiries into human faculty and its development*. London: Macmillan.

Gerlach, V. S., & Sullivan, H. J. (1967). *Constructing statements of outcomes*. Inglewood, CA: Southwest Laboratory for Educational Research and Development.

Ghiselli, E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology, 26*, 461–477.

Goldman, B. A., & Mitchell, D. F. (Eds.). (1990). *Directory of unpublished experimental mental measures* (Vol. 5). Dubuque, IA: Wm. C. Brown.

Goldman, B. A., & Osborne, W. L. (Eds.). (1985). *Directory of unpublished mental measures* (Vol. 4). New York: Human Sciences Press.

Green, B. F. (1992). Testing technology. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed.) (Vol. 4, pp. 1410–1414). New York: Macmillan.

Green, D. R. (1992). Test construction. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed.) (Vol. 4, pp. 1400–1406). New York: Macmillan.

Green, J. A. (1975). *Teacher-made tests* (2nd ed.). New York: Harper & Row.

Green, K. (1984). Effects of item characteristics on multiple-choice item difficulty. *Educational and Psychological Measurement, 44*, 551–561.

Guerra, E. L., Abramson, D. A., & Newmark, M. (1964). New York City foreign language oral ability rating scale. *Modern Language Journal, 48*, 486–489.

Hakstian, A. R., & Kansup, W. (1975). Comparison of several methods of assessing partial knowledge in multiple-choice tests: Testing procedures. *Journal of Educational Measurement, 12*, 231–239.

Hammill, D. D., Brown, L., & Bryant, B. R. (1992). *A consumer's guide to tests in print* (2nd ed.). Austin, TX: Pro-Ed.

Harrington, R. G. (Ed.). (1986). *Testing adolescents*. Kansas City, MO: Test Corporation of America.

Harrow, A. J. (1972). *A taxonomy of the psychomotor domain: A guide for developing behavioral objectives*. New York: McKay.

Hartigan, J. A., & Wigdor, A. K. (1989). *Fairness in employment testing: Validity generalization, minority issues, and the General Aptitude Test Battery*. Washington, DC: National Academy Press.

Herbert, E. A. (1992). Portfolios invite reflection—from both students and staff. *Educational Leadership, 49*(8), 58–61.

Herman, J. L. (1994). Item writing techniques. In T. Husén & T. N. Postlethwaite (Eds.), *International encyclopedia of education* (2nd ed.) (Vol. 5, pp. 3061–3066). Tarrytown, NY: Elsevier.

Hoffman, B. (1962). *The tyranny of testing*. New York: Crowell-Collier.

Hogan, J., & Quigley, A. (1994). Effects of preparing for physical ability tests. *Public Personnel Management, 23*(1), 85–104.

Hopkins, K. D., & Bracht, G. H. (1975). Ten-year stability of verbal and nonverbal IQ scores. *American Educational Research Journal, 12*, 469–477.

Hughes, S. (1995). Review of Denver: II. In J. C. Conoley & J. C. Impara (Eds.), *The twelfth mental measurements yearbook* (pp. 263–265). Lincoln: Buros Institute of Mental Measurements of the University of Nebraska-Lincoln.

Hunter, J. E. (1983). *The dimensionality of the General Aptitude Tests Battery (GATB) and the dominance of general factors over specific factors in the prediction of job performance* (USES Test Research Report No. 44). Washington, DC: U.S. Department of Labor.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.

Intelligence and its measurement: A symposium. (1921). *Journal of Educational Psychology, 12*, 123–147, 195–216.

Jensen, A. R. (1982a). Reaction time and psychometric *g*. In H. J. Eysenck (Ed.), *A model for intelligence*. Berlin: Springer-Verlag.

Jensen, A. R. (1982b). The chronometry of intelligence. In R. J. Sternberg (Ed.), *Advances in research in intelligence* (Vol. 1). Hillsdale, NJ: Erlbaum.

Kaufman, A. S. (1978). Review of Columbia Mental Maturity Scale. In O. K. Buros (Ed.), *The eighth mental measurements yearbook* (Vol. 1, pp. 299–301). Highland Park, NJ: Gryphon Press.

Kaufman, A. S., & Kaufman, N. L. (1983a). *Kaufman Assessment Battery for Children: Administration and scoring manual*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (1983b). *Kaufman Assessment Battery for Children: Interpretive manual*. Circle Pines, MN: American Guidance Service.

Keyser, D. J., & Sweetland, R. C. (Eds.). (1984–1994). *Test critiques* (Vols. 1–10). Austin, TX: Pro-Ed.

Kiess, H. O. (1996). *Statistical concepts for the behavioral sciences* (2nd ed.). Needham Heights, MA: Allyn & Bacon.

Klimko, I. P. (1984). Item arrangement, cognitive entry characteristics, sex, and test anxiety as predictors of achievement examination performance. *Journal of Experimental Education, 52*, 214–219.

Knobloch, H., & Pasamanick, B. (Eds.). (1974). *Gesell and Amatruda's developmental diagnosis* (3rd ed.). New York: Harper & Row.

Knobloch, H., Stevens, F., & Malone, A. (1987). *Manual of developmental diagnosis: The administration and interpretation of the Revised Gesell and Amatruda Developmental and Neurologic Examination*. Houston, TX: Developmental Evaluation Materials.

Kramer, J. J., & Conoley, J. (1992). *The eleventh mental measurements yearbook*. Lincoln: Buros Institute of Mental Measurements of the University of Nebraska-Lincoln.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives: Handbook: II. The affective domain*. New York: McKay.

Krug, S. E. (1993). *Psychware sourcebook* (4th ed.). Champaign, IL: MetriTech.

Kubiszyn, T., & Borich, G. (1990). *Educational testing and measurement*. Glenview, IL: Scott, Foresman.

Lacks, P. (1984). *Bender-Gestalt screening for brain dysfunction*. New York: Wiley.

Larson, G. E. (1994). Armed Services Vocational Aptitude Battery. In R. J. Sternberg (Ed.), *Encyclopedia of human intelligence* (pp. 121–124). New York: Macmillan.

Levine, M. (1976). The academic achievement test: Its historical context and social functions. *American Psychologist, 31*, 228–238.

Levy, P., & Goldstein, H. (1984). *Tests in education: A book of critical reviews*. New York: Academic Press.

Linn, R. L. (1992). Achievement testing. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed., pp. 1–12). New York: Macmillan.

Little, E. B. (1962). Overcorrection for guessing in multiple-choice test scoring. *Journal of Educational Research, 55,* 245–252.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute, 22,* 719–748.

Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). Baltimore: William & Wilkins.

McClelland, D. C. (1961). *The achieving society.* Princeton, NJ: Van Nostrand.

McDaniel, E. (1994). *Understanding educational measurement.* Madison, WI: Brown & Benchmark.

McNemar, Q. (1964). Lost: Our intelligence? Why? *American Psychologist, 19,* 871–882.

Mislevy, R. J., & Stocking, M. L. (1989). A consumer's guide to LOGIST and BILOG. *Applied Psychological Measurement, 13,* 57–75.

Mountain, M., & Snow, W. (1993). Wisconsin Card Sorting Test as a measure of frontal pathology: A review. *Clinical Neuropsychologist, 7,* 108–118.

Murphy, K. R. (1984). Review of Armed Services Vocational Aptitude Battery. In D. Keyser & R. Sweetland (Eds.), *Test critiques* (Vol. 1, pp. 61–69). Kansas City, MO: Test Corporation of America.

Murphy, K. R., & Davidshofer, C. O. (1994). *Psychological testing: Principles and applications* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Murphy, L. L., Conoley, J. C., & Impara, J. C. (Eds.). (1994). *Tests in print: IV.* Lincoln: University of Nebraska and Buros Institute of Mental Measurements.

Nixon, J. E., & Jewett, A. E. (1980). *An introduction to physical education* (9th ed.). Philadelphia: Saunders.

Norušis, M. J. (1992a). *SPSS/PC+ base system user's guide version 5.0.* Chicago, IL: SPSS.

Norušis, M. J. (1992b). *SPSS/PC+ advanced statistics version 5.0.* Chicago, IL: SPSS.

Norušis, M. J. (1992c). *SPSS/PC+ professional statistics version 5.0.* Chicago, IL: SPSS.

Pagano, R. R. (1994). *Understanding statistics* (4th ed.). Minneapolis/St. Paul: West.

Parents in Action on Special Education (PASE) v. Joseph P. Hannon, No. 74C 3586 (N.D. Ill. 1980).

Paterson, D. G., Elliott, R. M., Anderson, L. D., Tooks, H. A., & Heidbreder, E. (1930). *Minnesota Mechanical Ability Tests.* Minneapolis: University of Minnesota Press.

Paulson, F. L., & Paulson, P. R. (1992). *The making of a portfolio.* Unpublished manuscript. (Available from authors, 6800 Gable Parkway, Portland, OR 97225)

Pellegrino, J. (1992). Intelligence and aptitude measurement. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (6th ed.) (Vol. 2, pp. 627–637). New York: Macmillan.

Pellegrino, J. (1994). Abilities and aptitudes. In T. Husén & T. N. Postlethwaite (Eds.), *The international encyclopedia of education* (2nd ed.) (Vol. 1, pp. 1–6). Tarrytown, NY: Elsevier.

Plake, B. S., Ansorge, C. J., Parker, C. S., & Lowry, S. R. (1982). Effects of item arrangement, knowledge of arrangement, test anxiety and sex on test performance. *Journal of Educational Measurement, 19,* 49–57.

Platt, J. R. (1961). On maximizing the information obtained from science examinations. *American Journal of Physics, 29,* 111–122.

Preston, R. C. (1964). Ability of students to identify correct responses before reading. *Journal of Educational Research, 58,* 181–183.

Pyle, W. H. (1913). *The examination of school children.* New York: Macmillan.

Raymond, M. R., Webb, L. C., & Houston, W. M. (1991). Correcting performance-rating errors in oral examinations. *Evaluation and the Health Professions, 14*(1), 100–122.

Reckase, M. D., Ackerman, T. A., & Carlson, J. E. (1988). Building a unidimensional test using multidimensional items. *Journal of Educational Measurement, 25,* 193–203.

Reid, F. (1977). Alternative scoring formula for multiple-choice and true-false tests. *Journal of Educational Research, 70,* 335–339.

Reitan, R. M. (1966). A research program on the psychological effects of brain lesions in human beings. In N. R. Ellis (Ed.), *International research in mental retardation* (Vol. 1, pp. 153–218). New York: Academic Press.

Resnick, D. (1982). History of educational testing. In A. K. Wigdor & W. R. Garner (Eds.), *Ability testing: Uses, consequences, and controversies, Part 2* (pp. 173–194). Washington, DC: National Academic Press.

Ross, C. C., & Stanley, J. C. (1954). *Measurement in today's schools* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.

Rowley, G. L. (1974). Which examinees are most favoured by the use of multiple-choice tests? *Journal of Educational Measurement, 11,* 15–23.

Safrit, M. J., & Wood, T. M. (1995). *Introduction to measurement in physical educational and exercise science* (3rd ed.). St. Louis: Mosby.

Sample, B. M. (1992). *Performance assessment: An international experiment* (Report No 22-Caep-06). Princeton, NJ: Educational Testing Service.

Sattler, J. M. (1988). *Assessment of children* (3rd ed.). San Diego: Sattler.

Savitz, F. R. (1985). Effects of easy examination questions placed at the beginning of science multi-choice examinations. *Journal of Instructional Psychology, 12,* 6–10.

Scheuneman, J. (1979). A new method for assessing bias in test items. *Journal of Educational Measurement, 16,* 143–152.

Schmidt, F. L. (1985). Review of Wonderlic Personnel Test. In J. V. Mitchell, Jr. (Ed.), *The ninth mental measurements yearbook* (Vol. 2, pp. 1755–1757). Lincoln: Buros Institute of Mental Measurements, University of Nebraska.

Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology, 62,* 529–540.

Schmidt, F. L., & Hunter, J. E. (1978). Moderator research and the law of small numbers. *Personnel Psychology, 31,* 215–232.

Schmidt, F. L., & Hunter, J. E. (1981). Employment testing: Old theories and new research findings. *American Psychologist, 36,* 1128–1137.

Schmidt, F. L., Hunter, J. E., & Pearlman, K. (1982). Progress in validity generalization: Comments on Callender and Osbury and further developments. *Journal of Applied Psychology, 67,* 835–845.

Schmidt, F. L., Hunter, J. E., Pearlman, K., Hirsch, H. R., Sackett, P. R., Schmitt, N., Tenopyr, M. L., Kehoe, I., & Zedeck, S. (1985). Forty questions about validity

generalizations and meta-analysis with commentaries. *Personnel Psychology, 37,* 407–422.

Schoenfeldt, L. F. (1985). Review of Wonderlic Personnel Test. In J. V. Mitchell, Jr. (Ed.), *The ninth mental measurements yearbook* (Vol. 2, pp. 1757–1758). Lincoln: Buros Institute of Mental Measurements, University of Nebraska.

Shaha, S. H. (1984). Matching tests: Reduced anxiety and increased test effectiveness. *Educational and Psychological Measurement, 44,* 869–881.

Shavelson, R. J. (1996). *Statistical reasoning for the behavioral sciences* (3rd ed.). Needham Heights, MA: Allyn & Bacon.

Shea, C. (1994, September 7). "Gender gap" on examinations shrank again this year. *The Chronicle of Higher Education,* p. A54.

Simpson, E. J. (1966). The classification of educational objectives, psychomotor domain. *Illinois Teacher of Home Economics, 10,* 110–114.

Snow, R. E. (1992). Aptitude theory: Yesterday, today, and tomorrow. *Educational Psychologist, 27*(1), 5–32.

Sperry, R. W. (1968). Hemisphere deconnection and unity in conscious awareness. *American Psychologist, 23,* 723–733.

Stalnaker, J. L. (1951). The essay type of examination. In E. Lindquist (Ed.), *Educational measurement* (pp. 495–530). Washington, DC: American Council on Education.

Starch, D., & Elliott, E. C. (1913). Reliability of grading work in mathematics. *School Review, 21,* 254–259.

Stiggins, R. J. (1994). *Student-centered classroom assessment.* Columbus, OH: Macmillan.

Stoloff, M. L., & Couch, J. V. (Eds.). (1992). *Computer use in psychology: A directory of software* (3rd ed.). Washington, DC: American Psychological Association.

Strand, B. N., & Wilson, R. (1993). *Assessing sport skills.* Champaign, IL: Human Kinetics.

Sweetland, R. C., & Keyser, D. J. (Eds.). (1991). *Tests* (3rd ed.). Austin, TX: Pro-Ed.

Swiercinsky, D. P. (Ed.). (1985). *Testing adults.* Kansas City, MO: Test Corporation of America.

Thissen, D., & Steinberg, L. (1986). A taxonomy of item response models. *Psychometrika, 51,* 567–577.

Thorndike, R. L. (1985). The central role of general ability in prediction. *Multivariate Behavioral Research, 20,* 241–254.

Thorndike, R. L., & Hagen, E. P. (1959). *Ten thousand careers.* New York: Wiley.

U.S. Department of Defense. (1993). *ASVAB 18/19 technical manual.* North Chicago, IL: U.S. Military Entrance Processing Command.

U.S. Department of Defense. (1994). *Technical manual for the ASVAB 18/19 career exploration programs.* North Chicago, IL: U.S. Military Entrance Processing Command.

Vale, C. D. (1985). ASCAL: Item parameter estimation program [Computer program]. St. Paul, MN: Assessment Systems.

Vermont State Board of Education. (1991). *Looking beyond "the answer": Vermont's mathematics portfolio assessment program, pilot year, 1990–1991.* Montpelier: Vermont State Board of Education.

Vernon, P. E. (1960). *The structure of human abilities* (Rev. ed.). London: Methuen.

Vygotsky, L. S. (1962). *Thought and language* (E. Hanfmann & G. Vakar, Trans.). Cambridge, MA: MIT Press. (Original work published 1934)

Vygotsky, L. S. (1978). In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds. & Trans.), *Mind in society: The development of higher psychological processes.* Cambridge: Harvard University Press.

Weaver, S. J. (Ed.). (1984). *Testing children*. Kansas City, MO: Test Corporation of America.

Welsh, J. R., Kucinkas, S. K., & Curran, L. T. (1990). *Armed Services Vocational Aptitude Battery (ASVAB): Integrative review of validity studies.* San Antonio, TX: Brooks Air Force Base.

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design* (3rd ed.). New York: McGraw-Hill.

Wissler, C. (1901). The correlation of mental and physical tests. *Psychological Review Monograph Supplement, 3*(No. 6).

Wolf, D. P. (1989). Portfolio assessment: Sampling student work. *Educational Leadership, 46*(7), 35–36.

Woo, E. (1997, May 7). Few signs found of academic gender gap. *Los Angeles Times*, pp. A1, A33.

Wright, B. D., & Linacre, M. (1991). *A user's guide to BIGSTEPS.* Chicago: MESA Press.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis.* Chicago: MESA Press.

Yerkes, R. M. (1919). Report of the Psychology Committee of the National Research Council. *Psychological Review, 26*, 83–149.

Yerkes, R. M. (Ed.). (1921). Psychological examining in the United States army. *Memoirs of the National Academy of Sciences, 15.*

Zeidner, J., & Johnson, C. (1992). *Is personnel classification a concept whose time has passed?* Presented at the ARI Conference on Selection and Classification in the Military, Alexandria, VA.

الصفحة الأخيرة قبل الغلاف

للمعلومات المتعلقة بالقرص المرن المرفق، يمكن الرجوع إلى الملحق أ.

ملاحظة للزبون: يرجى قراءة التعليمات الآتية الخاصة بقوانين البرمجيات قبل الاستعمال:

إن هذا البرنامج يحتوي على ملفات لمساعدتك في استخدام النماذج الموصوفة في الكتاب المصاحب، وبفتح واستعمال رزمة البرامج فإن ذلك يعني أنك توافق على التزامك بالاتفاقية الآتية:

إن هذا المنتج البرامجي محمي بحقوق الطبع والنشر وإن جميع الحقوق محفوظة للمؤلف، أو جون وايلي أو المرخصين من قبلهم، وكزبون فإنك مرخّص لاستعمال هذه البرامج على حاسب آلي منفرد، إن نسخ هذه البرامج في وسط آخر أو صورة أخرى للاستعمال على حاسب آلي منفرد لا يخرق قوانين حقوق النشر الأمريكية، إن نسخ البرامج لشخص آخر يعد خرقاً لقوانين حقوق النشر الأمريكية.

لقد تم بيع هذه الرزمة من البرامج كما هي دون ضمان من أي نوع، سواء أكان صريحاً أو ضمنياً بما في ذلك، ولكن ليس محدداً به، الضمانات الضمنية للاتجار والملاءمة لغرض معين، ولا تتولى شركة وايلي ولا أي من عملائها أو موزعيها المسؤولية عن أي ملف مزعوم أو حقيقي ناتج عن استخدام أو عدم استخدام هذه البرامج. (لا تسمح بعض الولايات بعملية استثناء للضمان الضمني، ولذلك فإن الاستثناء ربما لا ينطبق عليك).


9789960541495